DOCUMENT RESUME

ED 437 629 CS 013 842

AUTHOR	Readence, John E., Ed.; Barone, Diane M., Ed.
TITLE	Revisiting the First-Grade Studies.
INSTITUTION	International Reading Association, Newark, DE.
PUB DATE	1998-00-00
NOTE	111p.; Reprinted from the Oct/Nov/Dec 1997 Themed Issue of "Reading Research Quarterly," v32 n4.
AVAILABLE FROM	Order Department, International Reading Association, 800 Barksdale Road, P.O. Box 8139, Newark, DE 19714-8139 ($7.96 members, $9.95 nonmembers). Tel: 800-336-7323, ext. 266 (Toll Free); Web site: http://newbookstore.reading.org.
PUB TYPE	Collected Works - Serials (022)
JOURNAL CIT	Reading Research Quarterly; v32 n4, Oct/Nov/Dec 1997
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	*Beginning Reading; Classroom Techniques; *Grade 1; Instructional Effectiveness; Primary Education; *Reading Instruction; *Reading Research
IDENTIFIERS	First Grade Reading Group Study; Historical Materials; Interpretive Research

ABSTRACT

This collection of articles commemorates the 30th anniversary of an influential study by Guy Bond and Robert Dykstrain reading research, known as the First-Grade Studies. These articles provide a view of the historical implications of the First-Grade Studies and how the Studies continue to provide valuable insights for those who seek answers about effective beginning reading instruction. The articles in the collection are entitled: "Introduction: Revisiting the First-Grade Studies: The Importance of Literary History" (John E. Readence and Diane M. Barone); "Contextualizing the First-Grade Studies: What Is the Best Way To Teach Children To Read?" (Michael F. Graves and Robert Dykstra); "The Cooperative Research Program in First-Grade Reading Instruction" (Guy L. Bond and Robert Dykstra); "The First-Grade Studies: A Personal Reflection" (P. David Pearson); "Connecting the Past with the Present: The Legacy and Spirit of the First-Grade Studies" (Lyndon W. Searfoss); and "Expanding the Boundaries: A Reaction to the First-Grade Studies" (Arlette Ingram Willis and Violet J. Harris). (NKA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 437 629

READING
RESEARCH
QUARTERLY

Revisiting the First-Grade Studies

JOHN E. READENCE
DIANE M. BARONE
Editors

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

J. Irwin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

CS 013 842

READING
RESEARCH
QUARTERLY

Reprinted from the
Oct/Nov/Dec 1997
the issue of
Reading Research
Quarterly
Vol. 32, No. 4
ISSN 0034-0553

Revisiting the First-Grade Studies

JOHN E. READENCE
University of Nevada, Las Vegas
Las Vegas, Nevada, USA

DIANE M. BARONE
University of Nevada, Reno
Reno, Nevada, USA

Editors

INTERNATIONAL
Reading Association
800 Barksdale Road, PO Box 8139
Newark, DE 19714-8139, USA
www.reading.org

BEST COPY AVAILABLE

IRA BOARD OF DIRECTORS

John J. Pikulski, University of Delaware, Newark, Delaware, *President* • Kathryn A. Ransom, Illinois State University, Springfield, Illinois, *President-Elect* • Carol Minnick Santa, School District #5, Kalispell, Montana, *Vice President* • Richard L. Allington, University at Albany–SUNY, Albany, New York • Betsy M. Baker, Columbia Public Schools, Columbia, Missouri • James F. Baumann, Department of Reading Education, University of Georgia, Athens, Georgia • James V. Hoffman, The University of Texas–Austin, Austin, Texas • Kathleen Stumpf Jongsma, Northside Independent School District, San Antonio, Texas • Adria F. Klein, California State University, San Bernardino, California • Diane L. Larson, Owatonna Public Schools, Owatonna, Minnesota • John W. Logan, Northbrook School District 27, Northbrook, Illinois • Lesley M. Morrow, Rutgers University, New Brunswick, New Jersey • Alan E. Farstrup, Executive Director

The International Reading Association attempts, through its publications, to provide a forum for a wide spectrum of opinions on reading. This policy permits divergent viewpoints without implying the endorsement of the Association.

Director of Publications Joan M. Irwin
Assistant Director of Publications Wendy Lapham Russ
Managing Editor, Journals Kate Tyler Wall
Associate Editor, Journals Rebecca Stewart
Assistant Editor, Journals June Hollins
Editorial Coordinator Cynthia Sawaya
Managing Editor, Books and Electronic Publications Christian A. Kempers
Associate Editor, Books Matthew W. Baker
Assistant Editor, Books Janet S. Parrack
Assistant Editor, Books Mara P. Gorman
Publications Coordinator Beth Doughty
Association Editor David K. Roberts
Production Department Manager Iona Sauscermen
Graphic Design Coordinator Boni Nash
Electronic Publishing Supervisor Wendy A. Mazur
Electronic Publishing Specialist Anette Schütz-Ruff
Electronic Publishing Specialist Cheryl J. Strum
Electronic Publishing Assistant Peggy Mason

Copyright 1998 by the International Reading Association, Inc.
All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, or any informational storage and retrieval system, without permission from the publisher.

CONTENTS

Introduction: Revisiting the First-Grade Studies: The importance of literacy history 1
JOHN E. READENCE
DIANE M. BARONE

Contextualizing the First-Grade Studies: What is the best way to teach children to read? 3
MICHAEL F. GRAVES
ROBERT DYKSTRA

The cooperative research program in first-grade reading instruction 9
GUY L. BOND
ROBERT DYKSTRA

The First-Grade Studies: A personal reflection 87
P. DAVID PEARSON

Connecting the past with the present: The legacy and spirit of the First-Grade Studies 92
LYNDON W. SEARFOSS

Expanding the boundaries: A reaction to the First-Grade Studies 98
ARLETTE INGRAM WILLIS
VIOLET J. HARRIS

AUTHORS

MICHAEL F. GRAVES is a professor of literacy education at the University of Minnesota. His research and development interests are in vocabulary learning and instruction, comprehension instruction, and effective teaching. He can be contacted at the Department of Curriculum and Instruction, 159 Pillsbury Drive SE, Minneapolis, MN 55455, USA, or by e-mail at grave002@maroon.tc.umn.edu.

ROBERT DYKSTRA is a professor of education Emeritus at the University of Minnesota. He assisted Guy Bond in coordinating the First-Grade Studies out of the University of Minnesota Coordinating Center and after Bond's retirement carried on the Center's work through its completion. He can be contacted at 1998 16th Street NW, New Brighton, MN 55112, USA.

LYNDON W. SEARFOSS is a professor of reading and school library science in the Division of Curriculum and Instruction at Arizona State University, where he teaches courses in reading methods, assessment, and current research issues in literacy. His recent interests include early intervention, emerging literacy, and classroom environments. He can be contacted at Reading/School Library Science, Arizona State University, Tempe, AZ 85287-0311, USA, or by e-mail at lyndon@IMAP2.asu.edu or Lnsear@aol.com.

P. DAVID PEARSON holds the John A. Hannah Distinguished Professorship of Education in the College of Education at Michigan State University, where he is a member of the Department of Teacher Education and the Department of Counseling, Educational Psychology, and Special Education. He continues to pursue a line of research related to reading instruction and reading assessment policies and practices at local, state, and national levels. His recent work focuses on attempting to validate standards-based approaches to portfolio and performance assessment. He can be contacted at 516 Erickson Hall, Michigan State University, East Lansing, MI 48224, USA.

ARLETTE INGRAM WILLIS received her Ph.D. from the Ohio State University. She is currently an assistant professor at the University of Illinois at Urbana-Champaign. Her research interests include the history of reading research in the United States, sociohistorical foundations of literacy, and teaching/learning multicultural literature for Grades 6–12. She can be contacted at University of Illinois at Urbana-Champaign, College of Education, Department of Curriculum and Instruction, 307 Education Building, 1310 South Sixth Street, Champaign, IL 61820-6990, USA, or by e-mail at aiwillis@uiuc.edu.

VIOLET J. HARRIS received her Ph.D. from the University of Georgia. She is currently an associate professor at the University of Illinois at Urbana-Champaign. Her research interests include children's literature, multiethnic children's literature, and literacy materials created for African American children prior to 1950. She can be contacted at University of Illinois at Urbana-Champaign, College of Education, Department of Curriculum and Instruction, 307 Education Building, 1310 South Sixth Street, Champaign, IL 61820-6990, USA, or by e-mail at vjharris@uiuc.edu.

INTRODUCTION

Reprinted from Reading Research Quarterly
Vol. 32, No. 4
October November December 1997
©International Reading Association

John E. Readence
Diane M. Barone

Revisiting the First-Grade Studies: The importance of literacy history

In this issue we boldly go beyond the tradition of research journals such as *Reading Research Quarterly* by reprinting "The Cooperative Research Program in First-Grade Reading Instruction," commonly known as the First-Grade Studies, written by Guy L. Bond and Robert Dykstra 30 years ago. The reprint of this study is introduced by Mike Graves and Bob Dykstra. Mike is at the University of Minnesota and was a colleague of Bob's before he retired. Bob, of course, is a coauthor of the First-Grade Studies. They provide the context for when and why the studies were conducted. Their prologue gives us a sense of the collaborative nature of these studies and provides a framework as we read or reread this classic study.

After reading the study, there are three retrospectives. The first retrospective, "Connecting the Past with the Present: The Legacy and Spirit of the First-Grade Studies," is written by Lyn Searfoss. Lyn was personally involved in the First-Grade Studies through his role as a graduate student of William D. Sheldon of Syracuse University, who directed an individual study in this project. Following Lyn's retrospective, Dave Pearson writes about "The First-Grade Studies: A Personal Reflection." Dave was a doctoral student at the University of Minnesota when Bond and Dykstra were completing the final stages of data analysis and manuscript development. Finally, Arlette Willis and Violet Harris conclude the retrospectives with "Expanding the Boundaries: A Reaction to the First-Grade Studies." We asked them to do this because we knew they would present another perspective by considering marginalized students who were not explicitly dealt with when the studies were reported.

At this point, you may be thinking: Why are we moving beyond the traditions of *RRQ* and perhaps opening ourselves up to criticism? There are numerous reasons why we have chosen to reprint the First-Grade Studies embedded in the context of a prologue and retrospectives:

- **The importance of history to literacy researchers**. Quoting from a recent Conversations article in *RRQ* written by Moore, Monaghan, and Hartman (1997): "History is a marginalized research genre among literacy professionals.... Unlike many disciplines, literacy education researchers have positioned history on the fringes" (p. 90). Additionally, IRA has "committed itself to the inclusion and encouragement of history as part of its ongoing activities" (*History of Reading News*, 1997). In response to these thoughts, we are being proactive in recognizing the value of literacy history by reprinting the First-Grade Studies, a classic research article.
- **The interest in beginning literacy**. A recent request for proposals has indicated that a new reading research center will have as its mission issues of beginning literacy. With this current attention to beginning literacy, it makes sense to revisit the First-Grade Studies, a research report that focused on the same issues.
- **This research report encourages methodological eclecticism**. The controversy between whole language and phonics instruction advocates has polarized beginning reading instruction. As Shanahan and Neuman (1997) have pointed out:

> eclecticism has been under renewed attack recently. Increasingly rancorous arguments about the superiority of various instructional approaches are evident, and some members of the research community have again become involved in the search for best method. An enduring legacy of this study has been a greater focus on teacher and learning situation characteristics rather than on methods and materials. (p. 208)

Interestingly, no study has challenged the basic findings of the First-Grade Studies.

• **Practically speaking, the First-Grade Studies are inaccessible**. The First-Grade Studies are no longer in print. Additionally, the research report comprises the entire issue of *RRQ*, so making a copy of the report is problematic. In fact, making a copy of the entire report violates copyright laws. Reprinting the research report remedies this ethical and legal dilemma and makes the First-Grade Studies accessible.

As Editors, we see that reprinting a research report, in this case the First-Grade Studies, provides a service to the literacy field. How many of our literacy colleagues have actually had formal exposure to literacy history through discussions centered on such reports? How many of our doctoral students take a course in literacy history? It would seem that in many cases, historical exploration has to be done individually, as a personal pursuit. The reprinting of this report provides a vehicle to begin solving this problem.

We conclude this Editorial by quoting from several of our Editorial Advisory Board members on the topic of reprinting classic studies.

"There is value at looking at what shaped our past. It's rather sad that we tend not to do so in a proper manner. However, that said...might it not make sense to do so in a manner where several individuals provide both current and historical analysis of the piece in question?" (N. Stahl, personal communication, March 3, 1997)

"I too think that there is a way to make use of our past in the present pages of *RRQ*. How best to do that is a matter for further discussion...but I think there must be ways to do it so that it animates our current discourse." (D.K. Hartman, personal communication, March 3, 1997)

"It's a shame for us to lose sight of our past. The idea of situating classic studies/articles within the context of current commentary seems intriguingly interesting to me." (M.R. Ruddell, personal communication, March 4, 1997)

It is with this spirit that we made the decision to reprint the First-Grade Studies. We hope the literacy educators of today and those who will move into these positions in the future enjoy and benefit from this venture. We encourage you to think about this issue of *RRQ* and the notion of reprinting classic articles, and let us know your thoughts through a Letter to the Editors.

REFERENCES

IRA invites task force on history. (1997, Spring). *History of Reading News, 20*(2).

MOORE, D.W., MONAGHAN, E.J., & HARTMAN, D.K. (1997). Values of literacy history. *Reading Research Quarterly, 32*, 90–102.

SHANAHAN, T., & NEUMAN, S.B. (1997). Literacy research that makes a difference. *Reading Research Quarterly, 32*, 202–210.

Reprinted from Reading Research Quarterly
Vol. 32, No. 4
October November December 1997

©International Reading Association

Michael F. Graves
Robert Dykstra

Contextualizing the First-Grade Studies: What is the best way to teach children to read?

The pendulum, a fitting if unfortunate symbol of our vacillating approach to beginning reading instruction, has a long history; and the fervor and rancor that have so often powered the reading pendulum can be found at various points in time. In the late 1700s, Samuel Heinicke, founder of the first German institute for training the deaf, expressed his disdain for the then prevalent spelling method of teaching reading with these words:

> The spelling method is a greater prejudice than burning witches and heretics; indeed it is a greater crime than the rack and all the inhumanities lumped together. For it is in defiance of all natural and revealed laws; it begets stupidity, illness, and death itself. It is child torture—a slower and surer child-murder. (quoted in Mathews, 1966, p. 35)

And in 1842, Horace Mann, Secretary of the Massachusetts State Board of Education, was only somewhat less dramatic in denouncing the alphabet-spelling method then prominent in the United States:

> Compare the above method [an approach Mann had observed in Prussia that began with meaningful words and only then analyzed the words into component parts] with that of calling up a class of abecedarians—or, what is more common, a single child—and while the teacher holds a book or card before him, with a pointer in his hand, says, *a*, and he echoes *a*; then *b*, and he echoes *b*, and so on until the vertical row of lifeless and ill-favored characters is completed, and then of remanding him to his seat, to sit still and look at vacancy. (reprinted in Smith, 1965, p. 78)

In the decades following Mann's pronouncement, the reading pendulum swung back and forth, and a variety of methods were used in American schools. But in the early 1900s, reading authorities began to reach more agreement on methods, and from about 1930 to the mid-1950s, the reading pendulum hung nearly motionless. The basal reader approach was firmly entrenched as the primary method of teaching reading. In response to a national survey of instructional practices in reading conducted by Ralph Staiger (1958), 69% of the respondents reported using one basal series, 20% reported using two basals, and 11% reported using three or more basals. That, as you can see, adds up to 100% of the respondents reporting that they used basals. The basals, of course, differed somewhat from each other, changed in some ways over the period, and included a number of different features. However, the feature most relevant here, and a common feature of most basals of that day, was their laid-back approach to phonics instruction. Typically, phonics teaching was delayed until after beginning readers had learned to identify about 50 sight words. Moreover, phonics was typically taught as a back-up word identification strategy, one to be used only after meaning clues and word structure analysis were unsuccessful.

This period of agreement came to an abrupt end, the pendulum was again put into motion, and the conflict that would prompt the First-Grade Studies was renewed with the publication of Rudolf Flesch's *Why*

Johnny Can't Read in 1955. Flesch attacked the basal approach to reading instruction with a vengeance:

> Ever since 1500 B.C. people all over the world—wherever an alphabetic system of writing was used—learned how to read and write by the simple process of memorizing the sound of each letter in the alphabet.... Except, as I said before, twentieth-century Americans.... We have thrown 3500 years of civilization out the window. (Flesch, 1955, pp. 4-5)

Flesch's book was immensely popular. It stayed on the bestseller list for more than 30 weeks; it was serialized in newspapers throughout the U.S.; and, in fact, it is currently on the bookshelves at Barnes and Noble. In the words of Jeanne Chall,

> Flesch challenged—strongly, clearly, and polemically—the prevailing views on beginning reading instruction, which emphasized teaching children by a sight method. He advocated a return to a phonic approach (early teaching of correspondences between letters and sounds) as the best—no, the *only*—method to use in beginning instruction. He found support for this view in his interpretation of the existing reading research, particularly the research comparing sight and phonic methods. (Oddly enough this same body of research formed the basis for the prevailing methods, and proponents of those methods used it to defend themselves.) (Chall, 1967, pp. 3-4)

Another factor influential in prompting, or at least enabling, the First-Grade Studies was the launching of Sputnik I in October of 1957. For 23 days, Sputnik I circled the earth, mocking the U.S. space effort with its constant beep. Many Americans believed that the Russians had demonstrated a scientific superiority that bode poorly for the outcome of the Cold War. Responding to this perceived threat, in 1958 the U.S. Congress passed the National Defense Education Act, which provided very substantial funding to channel students into courses of study the government deemed important to national security. Although much of this money went to science and engineering, funds also went to less technical areas; and the general recognition of the importance of improving education made securing funding for research such as the First-Grade Studies easier.

Prompted by the furor that had followed Flesch's book and probably encouraged about the possibilities of funding following Sputnik and the National Defense Education Act, the National Conference on Research in English established a special committee on reading research. The committee met for the first time at Syracuse University in 1959, when, according to Chall, the controversy over methods had reached its most bitter point, and laid the seeds for the First-Grade Studies.

At this meeting, the committee—composed of Guy Bond, Jeanne Chall, Theodore Clymer, Donald Durrell, William Sheldon, James Soffetti (a linguist at Syracuse), Ralph Staiger, and Russell Stauffer (chair)—agreed that the available research evidence, or at least the synthesis of the research that was currently available, was so vague, contradictory, and incomplete as to encourage conflicting interpretations. They further agreed on two complementary research programs. Chall decided to undertake a critical, systematic analysis of the research already available, a review eventually published as *Learning to Read: The Great Debate* (1967). The group as a whole decided to begin a large-scale cooperative experiment with clearly defined controls that could provide solid "evidence on whether some approaches were indeed more effective than others for specific outcomes in reading, for particular kinds of children, with particular kinds of teachers, and in particular kinds of school situations" (Chall, 1967, p. 5).

Subsequent to this meeting, an expanded group met at the University of Chicago in 1960 to establish the actual guidelines for a large-scale cooperative research program. In 1963, the Cooperative Research Branch of the U.S. Office of Education indicated its willingness to provide financial support for such a study and invited proposals. In 1964, Guy Bond was selected to orchestrate the research and synthesize and analyze the results of the various studies that would be completed as part of the program, and the University of Minnesota became the Coordinating Center for the research. In that same year, 76 proposals were submitted to the U.S. Office of Education, 27 of which were selected for funding. Then, in September 1964, the study itself began. It was a truly vast undertaking. Not only were 27 individual projects involved; the study quite literally spanned the length and breadth of the U.S., with, for example, Olive Niles directing a study in Massachusetts, George Spache directing one in Florida, and John Manning directing one in California.

As described in the 1967 *RRQ* report, the Studies were designed to gather data relevant to three basic questions:

> 1. To what extent are various pupil, teacher, class, school, and community characteristics related to pupil achievement in first-grade reading and spelling?
> 2. Which of the many approaches to initial reading instruction produces superior reading and spelling at the end of the first grade?
> 3. Is any program uniquely effective or ineffective for pupils with high or low readiness for reading? (Bond & Dykstra, 1967, p. 5 in original text/p. 9 in this book)

In doing so, the Studies investigated six instructional approaches: Basal, Basal plus Phonics, i.t.a., Linguistic, Language Experience, and Phonic/Linguistic.

Up to this point, we have attempted to describe briefly the genesis of the First-Grade Studies and the context in which they were undertaken. In concluding this introduction to the Studies, three points seem particularly worthy of comment.

First, with the exception of the language-experience approach, the essence of the other approaches and the differences among them had to do with the manner in which the code of written English was introduced to the beginning reader and the timing and relative emphasis given to instruction in the code. Today, of course, a great deal of the controversy over instructional methods revolves around these same issues, a situation that recently prompted the International Reading Association Board of Directors to issue a position statement on *The Role of Phonics in Reading Instruction* ("IRA Takes Stand," 1997).

Our second point is that the Studies represent a unique spirit of cooperation and selflessness among the reading education authorities of that era, most of whom were involved in the planning or the implementation of the research. These professionals, many of whom were authors of basal reading series, and many of whom had published the authoritative reading methods texts of the day, literally put their credibility on the line in this cooperative research endeavor.

Our third point has to do with the research zeitgeist at the time of the Studies. The Studies were planned and carried out shortly after the publication of Donald Campbell and Julian Stanley's classic *Experimental and Quasi-Experimental Designs for Research on Teaching* (1963) and well before Richard Snow's "Representative and Quasi-Representative Designs for Research on Teaching" (1974), certainly one of the first widely circulated educational articles to question the traditional approach to research. The studies are thus firmly in the Campbell and Stanley tradition and must be understood as a part of that tradition. This tradition has its limits, which have been prominently discussed in recent years, as well as its strengths, which have received much less attention recently.

The First-Grade Studies did not, of course, answer the question posed in our title. We now know, almost certainly better than we did 30 years ago, that the question has no simple answer, and is itself too simple. Still, it is in many ways very close to the questions that many of us in reading education hope to answer. If we now know that the sort of certainty sought from the Studies is not possible, for many of us the belief that research can play a vital and at least somewhat direct role in improving instruction is a central motivator for doing research. The First-Grade Studies provide valuable insights for those who seek to ask and answer questions about effective beginning reading instruction and for those who seek to answer other questions about reading instruction.

On the flyleaf of Alberto Manguel's *A History of Reading* (1996) appears this poetic description of learning to read:

> At one magical instant in your early childhood, the page of a book—that string of confused, alien ciphers—shivered into meaning. Words spoke to you, gave up their secrets; at that moment, whole universes opened. You became, irrevocably, a reader.

It is a wonderful piece of prose—sonorous, inspiring, eloquent. But few children have such an epiphany. For those many children for whom words do not shiver into meaning in some magical moment, research that probes questions like those investigated in the First-Grade Studies continues to offer the hope of opening universes.

REFERENCES

BOND, G.L., & DYKSTRA, R. (1967). The cooperative research program in first-grade-reading instruction. *Reading Research Quarterly, 2,* 1–142.

CAMPBELL, D.T., & STANLEY, J.C. (1963). *Experimental and quasi-experimental designs for research on teaching.* Chicago: Rand McNally.

CHALL, J.S. (1967). *Learning to read: The great debate.* New York: McGraw-Hill.

FLESCH, R. (1955). *Why Johnny can't read—and what you can do about it.* New York: Harper & Row.

IRA takes stand on phonics. (1997, April/May). *Reading Today,* pp. 1, 4.

MANGUEL, A. (1996). *A history of reading.* New York: Viking.

MATHEWS, M.M. (1966). *Teaching to read: Historically considered.* Chicago: University of Chicago Press.

SMITH, N.B. (1965). *American reading instruction.* Newark, DE: International Reading Association.

SNOW, R.E. (1974). Representative and quasi-representative designs for research on teaching. *Review of Educational Research, 44,* 265–291.

STAIGER, R.C. (1958). How are basal readers used? *Elementary English, 35,* 46–49.

Vol. 2, No. 4, 1967

In this issue

One of the truly unique and outstanding ventures in reading research is reported in this issue of the Quarterly. The contents of the issue provide in large measure the details, the procedures, and the major conclusions reached by Drs. Bond and Dykstra in the analysis of data which came to the Minnesota Coordinating Center from the 27 first-grade reading projects.

Most readers know that the Coordinating Center and the 27 individual projects in first-grade reading instructional methods were funded by the United States Office of Education. The Coordinating Center came into existence after many experimenters had made commitments to various experimental details and procedures and it was initiated through the cooperative effort of the 27 project directors.

The Coordinating Center was instrumental in developing new generalizations about first-grade reading instruction. In addition, and certainly as important, the entire project demonstrates the possibilities, the rewards, and some of the problems of truly cooperative research in reading.

Professor Donald D. Durrell of Boston University deserves special recognition for his diligent efforts in gaining support for the project. Beginning with a meeting at Syracuse University in 1959, many members of the reading profession were interested in the concept of cooperative research in reading and worked to gain its acceptance. However, the success of Durrell in enlisting support within the USOE was one of the major factors in getting the study underway.

This issue of the Quarterly must be studied carefully and as a total document for its full value to be realized. A quick reading or a few quotes pulled from the summary sections will not do justice to the content of the report or the spirit of the inquiry.

Theodore Clymer, EDITOR
Edward G. Summers, ASSOCIATE EDITOR

The cooperative research program in first-grade reading instruction

GUY L. BOND & ROBERT DYKSTRA

Review of literature **10**
The initial teaching alphabet / Phonic methods / Individualized methods / Linguistic methods / Language experience methods / Sex differences and reading / Summary of review literature

Overview of Cooperative Research Program in First-Grade Reading Instruction **17**

Procedures **17**
Role of the Coordinating Center / Data collected / Common experimental guidelines / Organization of the data / General procedure of analysis

Analysis of relationships **21**
Relationships between readiness and reading / Intercorrelations among group and individual achievement measures / Relationships between teacher, pupil, class characteristics, and achievement

Evaluation of effectiveness of primary reading programs in the Cooperative Research Program **27**
General procedures / Description of analysis / An illustration of the analysis (Basal versus i.t.a) / Basal versus Basal plus Phonics comparisons / Basal versus Language Experience comparisons / Basal versus Linguistic comparisons / Basal versus Phonic/Linguistic comparisons / The practicality of significant differences / Variability within treatments / Summary of treatment comparisons / Limitations of the analysis of treatment comparisons

Analysis of treatment by readiness level **62**
Blocking on intelligence / Blocking on phonemes / Blocking on letter knowledge / Summary

Summary and conclusions **71**
Analysis of relationships / Analysis of methodology / Analysis of treatment by readiness level

Appendix I **77**
Individual studies of Cooperative Research Program in First-Grade Reading Instruction

Appendix II **83**
Instructional materials

Appendix III **84**
Tests

References **85**

AUTHORS

GUY L. BOND, Director of the Coordinating Center for the First-Grade Reading Studies, was formerly Professor of Education at the University of Minnesota, now retired. He received his advanced degrees from Columbia University, where he also served as a visiting professor during the summers of 1940 and 1951. During World War II, as a Lieutenant Commander, he was the officer-in-charge of the Selection and Classification Test Construction Unit of the U.S. Navy. He is the author of five well known professional books, *Teaching the Child to Read, Developmental Reading in the High School, The Diagnosis and Treatment of Learning Difficulties, Child Growth in Reading,* and *Reading Difficulties: Their Diagnosis and Correction* as well as several monographs in the field of reading and numerous professional articles. He is the co-author of the *Living Literature Series,* a set of literary readers for grades three to eight and senior author of *The Developmental Reading Series* for grades one to eight. He is also co-author of the *Developmental Science Series* for grades one to six. His experience has been broad and varied including classroom experience in both the elementary and secondary schools, practical experience with the New York State remedial reading program, and many years' experience in training teachers and supervisors. Dr. Bond has participated in workshops and reading conferences throughout the United States. He also has studied the problems of teaching reading, as well as the methods and materials used, in several foreign countries, namely, New Zealand, Australia, Scotland, England, Japan, Philippines, and Thailand.

ROBERT DYKSTRA, Associate Professor of Education at the University of Minnesota, received his Ph.D. in Educational Psychology from that same institution. His teaching and research interests are concerned primarily with elementary reading and language arts. Dr. Dykstra has taught in Wisconsin elementary schools and is currently serving as Director of the Coordinating Center for the follow-up studies in the USOE-supported Cooperative Research Program in First-Grade Reading

Reprinted from Reading Research Quarterly
Vol. 32, No. 4
October/November/December 1997
©International Reading Association

Guy L. Bond
Robert Dykstra

The cooperative research program in first-grade reading instruction

The present study was designed to obtain information relevant to three basic questions: (1) To what extent are various pupil, teacher, class, school, and community characteristics related to pupil achievement in first-grade reading and spelling? (2) Which of the many approaches to initial reading instruction produces superior reading and spelling achievement at the end of the first grade? (3) Is any program uniquely effective or ineffective for pupils with high or low readiness for reading?

Every year hundreds of thousands of children begin the complex task of learning to read. For most children, growth in reading is a successful undertaking. For many, however, the progress is slow, and for others learning to read appears to be an unobtainable accomplishment. There is a continuous search for new ways to teach reading which will prevent the difficulties these children encounter, thereby enabling all children to become successful readers. Moreover, even for those children who have apparent success in learning to read, there is always the question of whether or not a different approach would have enabled them to become even more mature and diversified readers.

In recent years, many new approaches to reading instruction have been suggested. Many questions have been raised about current methods of teaching reading. In fact, the teaching of beginning reading has been and continues to be a popular subject for debate among reading experts and the general public alike. Even though a great deal of research has been devoted to the problem, there are still a number of controversies concerning instructional procedures in beginning reading. While many new approaches to initial instruction have been formulated and implemented, they have not been subjected to comparative research to any extent. Furthermore, most of the research has been conducted in a piecemeal fashion by independent investigators. As a result, comparisons among the individual studies have been difficult for a number of reasons:

(1) Independent investigators have used different tests to measure reading readiness and reading achievement. Norming populations for the various tests may be quite different and, as a result, it is difficult to compare achievement of pupils whose reading ability has been assessed by different instruments.

(2) The extent to which investigators have assessed and/or controlled such factors as experiential background of children, class size, teacher competence, enthusiasm for the teaching method employed, and other such variables has varied from study to study.

(3) Research designs and methods of statistical analysis have varied from study to study.

(4) Evaluation of postinstructional reading ability has been incomplete and inappropriate.

(5) Experimental guidelines such as length of instructional period have varied considerably in independent investigations. Furthermore, the length of some experimental periods has been inadequate for demonstrating long-range effects of approaches to initial reading instruction.

(6) Methods, materials, and experimental populations have not been adequately described in order to make comparisons between studies possible.

The Cooperative Research Studies in First-Grade Reading Instruction were designed to overcome many of these difficulties. The unique contribution of this research program was its provision for coordination of a number of individual reading studies, thereby making possible the exploration of the relative effects on early reading growth of various approaches to initial reading instruction under similar experimental conditions.

The current study is the report of the Coordinating Center of the Cooperative Research Program in First-Grade Reading Instruction. The data used were made available by the 27 individual projects carried out under the Program.

Review of literature

A large number of studies have been reported which aim to determine the efficiency of different methods of teaching beginning reading. The results of these experiments have often been inconclusive and, at times, contradictory. Because of the variability of the results reported by investigators, the reader is left with little irrefutable evidence from which to determine the comparative efficiency of the methods of teaching reading which have been studied. After reviewing literature concerning the various methods of teaching reading, Gray (1960), in the reading section of the *Encyclopedia of Educational Research*, concluded that the issue was not which method was better, but rather what the contribution of each method is.

Russell and Fea (1963) in their chapter, "Research on Teaching Reading" in the *Handbook of Research on Teaching*, stated that historically

> thinking in the field has moved away somewhat from an either-or point of view about one method or set of books to a realization that different children learn in different ways, that the processes of learning to read and reading are more complex than we once thought, and that the issues in reading instruction are many sided. (p. 867)

Most of the studies reviewed in this paper involve comparative studies between basal reading programs and some other approach to teaching children to read.

The basal reading series has been the backbone of the elementary reading program in the United States for many years. In 1957, David Stewart reported on a questionnaire survey of practices in teaching reading. The survey included school systems in 107 cities of over 25,000 population, in 40 states. These districts were responsible for the reading instruction of 250,000 children. Stewart concluded that all schools were making use of one or more basic reading series. Austin and Morrison (1963) reported that, in more than 95% of the elementary schools they studied, the teachers relied heavily upon a basal or cobasal reading program.

The review of literature that follows is not all-inclusive, rather it is intended to be a sampling of the results reported and of the literature available in the areas of concern. In the literature concerning some approaches to beginning reading instruction, there is a great deal of opinion and intuitive writing available, some of which has been included in this review.

The initial teaching alphabet

One of the more recent innovations in the teaching of beginning reading is the Initial Teaching Alphabet, hereafter referred to as the i.t.a. i.t.a., developed by Sir James Pitman, was originally called the Augmented Roman Alphabet. According to Downing (1964, p. 15), the major investigator in i.t.a. studies in England, "the initial teaching alphabet has been designed for the specific purpose of helping children in the early stages of learning to read."

Using the i.t.a. materials, children learn to read using textbooks printed in a special alphabet consisting of 44 characters. Twenty-four of these are Roman or Latin characters used in traditional English print. There are 20 new letters, most of which are augmentations of the Roman alphabet. These characters are designed to regularize the coding of the basic sound units of English. Only lower case letters are used in order to reduce the number of characters necessary for the children to learn. Upper case letters are represented by larger forms of the i.t.a. lower case shapes. After pupils have gained confidence and fluency in reading i.t.a., they are to transfer their skills and confidence to reading material printed in the conventional characters.

Downing (1964) claimed that the teacher generally does not have to modify his teaching methods except for some adaptations forced upon him by the nature of i.t.a. In other words, i.t.a. may be used with a whole-word method, a phonics method, or a language experience method. Downing stated,

> Summing up, it is claimed that i.t.a. should help the global approaches to the teaching of reading because it makes the visual patterns invariable, and it should help the phonic approach, because in i.t.a. each symbol represents, with certain exceptions, one phoneme (p. 21).

Downing further asserted that the use of i.t.a. with primary-grade children would, in all likelihood, lead the pupils to learn that there is a systematic relationship between spelling and speech, and that experience with i.t.a. might help the children in their general intellectual development.

Downing (1963) stated there are two criticisms sometimes leveled at i.t.a.: (1) there may be too many characters for the children to learn, and (2) the new characters may be too difficult for the beginner to form with a pencil. Downing refuted the first of these claims by explaining that with traditional orthography, the children have to learn many more characters than by the i.t.a. system. The second criticism of i.t.a. was answered by Downing in a rather lengthy explanation of the system by which the children are trained to form the i.t.a. characters.

According to Downing (1963), the results of a longitudinal British study, begun in 1961, have shown that children using i.t.a. recognize more words in print, comprehend more continuous prose in print, read faster and more accurately, and progress through reading instruction more rapidly than children using the conventional type of basal reading program. Head teachers at the experimental English schools have reported that the i.t.a. medium appears to have raised the beginners' level of self-confidence, increased their enthusiasm for and interest in independent reading, allowed them to be more independent in their work, resulted in a marked improvement in creative writing, and permitted their thoughts to flow more naturally.

After one year of an i.t.a. study conducted in Bethlehem, Pennsylvania, Mazurkiewicz (1964) reported generally favorable conclusions in favor of his experimental (i.t.a.) group over his control group. However, in a discussion of the results of the same i.t.a. program after two years of the study in Bethlehem, Rebecca Stewart (1965) presented generally inconclusive results. At that time, there was no particular statistically significant advantage for either the i.t.a. groups or the groups which learned to read by means of traditional basal readers.

Chasnoff's (1965) study, in which the teacher variable was controlled, yielded scores for the total experimental group significantly higher in word reading, word study, and spelling, with respect to scores on the Stanford, Form W, when the experimental group was examined with a test transliterated into i.t.a. and the control group was examined by the same test in traditional orthography. On the Stanford, Form X, no significant differences were indicated with respect to scores gained on tests with all subjects taking the tests in traditional orthography. On a comparison of scores assigned to 616 writing samples gained from the total population, the differences of means for the experimental group was significantly higher at the .01 level. The scores generally appeared to give an advantage to the total experimental group, especially to subjects from three particular schools involved in the study, and to children who scored 35 to 44 on the California Test of Mental Maturity.

The results of many of the i.t.a. studies to date have been favorable for the i.t.a. groups. Most researchers recognize the need for follow-up studies on the effects of i.t.a.

Phonic methods

Some confusion exists as to the meanings of the terms *phonics* and *phonetics*. Often, these terms are used interchangeably in discussions of reading instruction practices. Phonics is a term for the practices of teaching reading in which individual letters of the alphabet are matched with the specific sounds of English pronunciation. Phonetics, on the other hand, is the process of systematic analysis and description of the vocal sounds, or phonetic features, of a language. It must be remembered, however, that the terms phonics, phonetics, or phonetic methods often refer to an entire method of teaching reading, supplementary teaching of phonics as an area of study in its own right, or the teaching of phonics as a part of another method.

Phonics can be further classified as either synthetic or analytic. The synthetic method is based upon the belief that the child should be taught certain letter-sound relationships of word elements before beginning to read and then be taught to synthesize word elements learned into whole words. Most older methods of teaching phonics were usually synthetic. The analytic method is based upon the belief that children should be taught whole words and then, through various analytic techniques, be taught to apply letter combinations learned in familiar words to sounding out new words.

There is no paucity of literature concerning the use of phonics in teaching beginning reading. In 1958, at the University of Pittsburgh, Morrone reviewed 198 references on phonics for a doctoral study. He suggested that no incontrovertible evidence was revealed by scientific investigations of phonics in reading and spelling. Morrone further stated:

> Disagreement exists as to the approach and amount of phonic instruction teachers should utilize in reading; however, most of the scientifically accurate experiments show that phonics has considerable value to the learner in the reading process (1958, p. 14).

Harrington and Durrell (1955) concluded that auditory and visual discrimination and phonic ability were more important than mental age for learning to read. Gates and Russell (1938) concluded that a program containing little or no phonetic analysis was not as good as one containing moderate amounts of informal word analysis. They also concluded that a moderate amount of informal word analysis made a better program than one which contained large amounts of drill-type phonics. In a study of the Carden method of teaching phonics, Gates (1961, p. 252) stated:

> The findings of this study do not suggest that teaching phonics is futile or unnecessary. They show merely that the much less complex and less rigid programs employed

> in most American schools during the past decade produce reading abilities equal to, or somewhat better, than the Carden system in much less time and with less effort.

Rudisill (1957), in a study designed to investigate the interrelations between phonic knowledge, reading achievement, spelling achievement, and mental age, found that a knowledge of phonics makes a substantial contribution to achievement in reading.

In a study designed to determine what relationships exist between phonic ability and reading ability, Tiffin and McKinnis (1940) tested 155 pupils in Grades 5, 6, 7, and 8 on the Iowa Silent Reading Test and the New Stanford Reading Test. An individual phonic test using nonsense words was also administered. The investigators concluded that phonic ability is significantly related to reading ability, and that a reading program should include direct or indirect instruction in the principles of phonics.

Few research studies have been reported condemning phonics. Two such studies were reviewed in a publication issued in 1963 by the University of the State of New York. The first of these studies was reported by Dumville in 1912. In his study, Dumville (University of the State of New York, 1963) used only 36 elementary school children about whom he reports no information concerning mental age, chronological age, sex, or any of the other background information usually considered necessary in such an experiment today. In Dumville's experiment, the children were divided into two groups, a phonics group and a look-and-say group.

> The look-and-say group was given a list of words in phonetic transcript and regular spelling and told to learn them as whole words. The phonics group was given a table of phonetic symbols, their sounds, and sample words in phonetic transcript. They were also given the same list of words in phonetic spelling and regular spelling. Both groups were given 15 minutes to learn the words, the former group learning whole words with the latter applying word analysis. They then had a practice test. The final tests were two extracts written in phonetic symbols; one contained the words on the list and the other was totally unfamiliar. Each student was tested individually for speed and number of mistakes. The results showed that the group using the look-and-say method was better on both tests in speed and lack of errors (University of the State of New York, 1963, pp. 5–6).

Obviously, the results of Dumville's experiment are somewhat questionable. The other study mentioned above was carried out by Mosher and Newhall and reported in 1930. This study, although better designed than Dumville's, is also open to question. The investigators concluded "that the differences were not significant enough to warrant spending time on phonics" (University of the State of New York, 1963, p. 7).

In a comparative study, Sparks and Fay (1957) concluded that, at the end of Grade 1, the *Phonetic Keys to Reading* method produced superior results in comprehension and vocabulary over a basal reading program. At the end of Grade 2, the phonic method led to superior results in comprehension only. However, at the end of Grade 4, no significant differences were found between the two groups in reading comprehension, vocabulary, or speed. At this time, the basic reading group was superior in reading accuracy. Sparks and Fay concluded that neither method was superior to the other.

In a study by Buswell (1922), an elaborate phonic method was contrasted with another method emphasizing thoughtful reading attitude and meaningful experience. He found that the phonic method promoted progress in the ability to follow the lines and pronounce the words, but it did not create a vital concern for the content. The method emphasizing thoughtful reading attitude and meaningful experience promoted a keen interest in the content, but slower progress was noted in word recognition and in the ability to follow lines.

McDowell (1953) compared five schools using a synthetic phonic approach with five schools using a basal reading approach where phonics were taught as a part of the word attack skills. Using matched pairs, McDowell tested the children on the Iowa Silent Reading Test and the Metropolitan Achievement Battery. On the Iowa test, the basal group obtained better scores on all measures except Directed Reading and Alphabetizing. Significant differences favoring the basal group were found on Word Meaning and in the medial reading scores. Significant differences favoring the phonics group were found in Alphabetizing. No significant differences were found on the Metropolitan tests in reading, vocabulary, and language. McDowell also compared pupils who had missed the first five months of phonics instruction with a matched group who had had the entire phonics program. Scores on the Iowa test showed no significant differences in the two groups. McDowell concluded that the phonics program was not accomplishing the results it claimed.

In a study of Philippine children, Tensuan and Davis (1963) compared a phonic method (called a cartilla method which involves learning grapheme-phoneme associations) with a combination method (a multiple approach similar to basal programs used in the United States). In the phonics approach, pupils were first taught the sounds of letters and diphthongs and, next, to identify sounds and words and to blend sounds. In the combination approach, interest in word knowledge was first aroused and whole words were associated with their meanings, after which letters and diphthongs were associated with the sounds and words that the pupil was already reading by sight. The expected difference was in favor of the cartilla method because there is a close correspondence between graphemes and phonemes in the Filipino language. No significant differences were found between the two groups on paragraph comprehension or language usage. The differences found, though not significant, favored the combination method.

Tate (1937), in a rather limited study, compared two groups matched in chronological age, mental age, and I.Q. One of the groups was taught phonics in 30-minute drill periods, while the other group had drill in word recognition and other skills. Both groups used an identical basic reading series for reading instruction. The results showed that the phonics group made greater gains in word recognition, while the other group gained more in word, phrase, and sentence reading, and in reading directions. From his data, Tate concluded that overemphasis upon phonics interferes with comprehension and that formal phonics drill is undesirable.

In a longitudinal study comparing synthetic and analytic approaches to teaching phonics, Bear (1964) found that, after 1 year of reading instruction, differences in performance on the Gates Primary Reading Tests and the Metropolitan Achievement Test favored the group using the synthetic method. A follow-up study of the pupils, after they had completed the sixth grade, found that the group which had utilized the synthetic method of phonics in the first grade was significantly superior in performance on the vocabulary section of the Gates Reading Survey, although no differences were found between the groups on the comprehension and speed sections of the test.

In another recent study, Bleismer and Yarborough (1965) concluded that the synthetic approach tended to be significantly more productive in terms of specific reading achievement than did the analytic approach.

Agnew (1939), working with children in Durham and Raleigh, North Carolina, used matched pairs to compare results of a program which stressed phonics (Durham) with a program which did not stress phonics (Raleigh). On Gates's four tests of phonetic ability, the Gates Word Pronunciation Test, Pressey Diagnostic Test–Vocabulary, Gray Oral Check Tests, and the Eye-Voice Span Test, the pupils from the program which stressed phonics were superior. On the Gates Silent Reading Tests, the groups were approximately equal, with a slight superiority of those in the stressed phonics program. The pupils from the stressed phonics program appeared to be slower in oral reading but more accurate. Agnew concluded that large amounts of phonetic training were valuable in facilitating skill development of the nature evaluated in the study.

In reviewing research on teaching reading, Russell and Fea (1963) concluded:

> The many "phonics versus whole-word" experiments in teaching have contained uncontrolled variables. Experiments designed to determine the relative effectiveness of different amounts of phonics, or the value of phonics at different maturation levels, have been more successful (1963, p. 875).

Dolch and Bloomster (1937) studied the correlation between phonics and mental ability. They concluded that the application of phonic principles required higher mental development than the memorization of sight words. Their results showed that children below the mental age of seven years made only chance scores on Tests 1 and 2 of the Basic Reading Tests, Word Attack Series, and concluded "as far as this experiment indicates, a mental age of seven years seems to be the lowest at which a child can be expected to use phonics, even in the simple situation provided by these two tests" (1937, p. 204).

On the other hand, Olson (1958), after testing first-grade children in September, November, and February and comparing results, concluded there is no support for the assumption that a mental age of seven is necessary for the use of phonics.

From the evidence reported concerning the use of phonics in teaching children to read, there can be little doubt that phonics should be an important part of the reading program. However, there is disagreement on the type of phonics approach which should be used and on the amount of phonics which should be included in the reading program. It seems apparent, from the studies reviewed, that phonics does not contribute much to children's comprehension of what is read; its main contribution is in the area of word recognition.

Linguistic methods

Recently, linguists have been attempting to apply their scientific knowledge of language to reading by suggesting linguistic generalizations which they believe are applicable to reading. Bloomfield and Barnhart (1961) developed a system in which children are systematically introduced to the written symbols that represent specific phonemes. Fries (1962) has developed an approach to teaching reading which he calls "linguistically sound." Fries stresses contrastive patterns of letters in words that function in consistent ways. He stated that a "structural base that constitutes the essential feature of every part of language" exists and that "...structuralism not only requires us to abandon our word-centered thinking about language; it demands that in every aspect of language we must shift from an *item-centered* view to one that is *structure-centered*" (1962, p. 64).

Strickland (1963) has interpreted linguistics to mean that: (1) the whole-word meaning approach without teaching the spoken linguistic forms symbolized by written shapes is wrong; (2) sounds are represented by letters and not letters by sounds; and (3) reading textbook writers need to give more attention to sentence structure, and systematic progress in sentence structure in sentence difficulty.

Sister Mary Fidelia (1959) compared two groups: one was taught by a linguistic approach, based on the work of Bloomfield; the other, by a phonics approach using a series of phonics workbooks called *Phonics We Use*. Both groups also used a basal reading series. No significant differences in reading achievement were found between the two groups.

Sister Mary Edward Dolan (1964) attempted to answer the question of whether introducing only regularly represented words in the early stages of reading is wise in light of the multitude of inconsistencies which the child encounters in later reading. She compared groups of fourth-grade pupils from parochial schools in Detroit, Michigan, and Dubuque, Iowa. One group used a composite basal method alone, while the other used a modified linguistic method in addition to a composite basal approach. In the modified linguistic approach, word recognition began with learning the alphabet and proceeded from sets of words and syllables with regular phoneme-grapheme relationships to more irregularly spelled words. The group using the linguistic material in addition to the basal program performed significantly better on a majority of the reading tests administered. The author did hypothesize, however, as to factors other than the modified linguistic material which may have been responsible for this superior achievement.

In a study on linguistics and reading, Goldberg and Rasmussen (1963) reported favorably on a "linguistic or phonemic-word" approach.

A number of studies have been done in attempts to discover a relation between reading and language structure. Gibbons (1941) noted that a close association existed between the reading level of children and their ability to understand the structure of sentences. MacKinnon (1959) found that beginning readers attempted to substitute syntactic patterns which they had previously read and with which they were familiar in place of unfamiliar patterns in attempting to decode new reading material.

In a comparative analysis of pupils' oral language patterns and the language patterns expressed in basal readers, Strickland (1963) concluded that pupils' language patterns are much more varied than patterns found in basal readers. She also reported that children who ranked high in silent reading comprehension made more use of common structural patterns, movables, and elements of subordination and elaboration than did children who ranked low on these variables.

Ruddell (1963) found that children's reading comprehension scores at the fourth-grade level are significantly higher on reading passages using only high frequency patterns of their oral language structure when compared to reading passages encompassing only low frequency patterns of their oral language structure.

Davis (1964) reported a comparative study involving a linguistic approach to first-grade reading instruction. Two of four groups used a basal reader program with a supplement of 111 daily lessons in linguistics. The other two groups spent an equal amount of time on only the basal reading program. The linguistic lessons involved 73 lessons in methods of word recognition applying phonemic-graphemic analysis, 20 lessons in identification of writing systems, 15 lessons on the alphabetical principle of writing, and 5 lessons on the structural patterns of written American alphabetical language. The investigator found significant differences in favor of the experimental groups on a battery of tests at the conclusion of the experiment.

It is apparent that there has not been a great deal of research evidence reported concerning the use of linguistic methods of teaching reading. The studies which have been done indicate there is value in the use of linguistic principles in designing a reading program. However, there is some indication that sentence structure should receive more consideration in the construction of reading materials. More research is needed concerning the use of linguistic principles in the teaching of reading.

Individualized methods

Rather than a single method, individualized reading programs are characterized by a multiple approach in which the teacher chooses his method or methods according to the child. No attempt is made to force predetermined standards upon the children and each child is to progress at his own rate. The classroom organization is such that the child receives more individual attention concerning his reading problems. Individualized reading does not mean complete elimination of group procedures. The choice between individual and group procedures is governed by the purposes of the reading being done at the time. The initial stages of the individualized approach are comprised of conversation, storytelling, reading aloud, and possibly an approach similar to basal reading approaches. As the children gain some fluency in reading, a variety of books chosen to suit a wide range of ability levels and interests is provided and, with the assistance of the teachers, the children choose from these according to their readiness, needs, and interests. Opinions of the value of the individualized reading programs are conflicting. There is not a great deal of research evidence available concerning this method.

Gates et al. (1926), in an early study, compared the relative merits of a systematic method and an opportunistic one in which the reading instruction was highly individualized. In respect to silent and oral reading, the investigators found that the results favored the systematic approach. However, the results also appeared to indicate that the method with highly individualized reading instruction was advantageous in respect to the development of interest, initiative, determination, and other personal and social traits.

Anderson et al. (1956b) compared a group using highly individualized methods with another using a systematic basal approach. The children in the individualized methods group were introduced to reading when they were ready for it and were permitted to choose the books that were read. Some use of basal readers was made with the individualized group, but they were not followed systematically. The individualized methods were used in a laboratory school where the average I.Q. was ten points higher than the public school group using the systematic basal approach. The investigators concluded that "the systematic approach employed by the public schools enables the children to learn to read early and reduces the individual variation in age of learning to read" (Anderson et al., 1956b, p. 107). The mentally superior group did not overtake the public school group until they were, on the average, 132 months of age.

In a study of primary reading instruction patterns, Sperry (1961) concluded that individualized reading classes showed significantly higher reading achievement than classes grouped by ability.

In a comparative study at the second-grade level, McCristy (1957) found the four classes in the experimental group (individualized method) were superior in total reading gains, vocabulary growth, and comprehension to the four classes in the control group (ability grouping).

In an investigation of individualized reading and the basal approach with primary children, Carline (1960) found no significant differences between the two approaches. Sartain (1960) compared the progress in reading skills of second graders taught by an individualized approach with those taught by a basal reader method. Significant differences were found between methods only for pupils of lower ability whose gains on word-recognition tests under the basal approach were superior to the gains made by lower ability pupils under an individualized approach. The methods were reversed after three months and the investigator found that significantly greater gains were made during the first three months of school, regardless of method employed.

Safford (1960) conducted a study of individualized reading involving seven classes in Grades 3 through 6. Results on the California Achievement Battery, administered at the end of the experimental period, showed the classes made gains considerably below national or district norms. Safford concluded that, for the majority of pupils in the classes involved, individualized reading resulted in lower gains, and that the use of self-selected reading methods achieved no significantly different results with pupils of high ability or those with average ability.

Zirbes et al. (1925) studied extensive individual reading instruction with short comprehension checks as compared with independent silent reading with second graders. The average growth in reading was about the same for both groups. The authors concluded that the more intelligent children profited more from the independent silent reading while the slower children profited more from intensive instruction.

Inconclusive and conflicting evidence has been reported concerning the use of individualized methods in beginning instruction in reading. There is some evidence to indicate that individualized methods may be more valuable for high ability children and that it may result in higher motivation and interest on the part of the children.

Language experience methods

The language experience approach to teaching beginning reading attempts to bring the communication or language arts skills (speaking, listening, writing, and reading) together as a unit. According to R.V. Allen (1964, p. 60), "The 'togetherness' of skill development makes possible the continuing use of each child's own experience background and thinking as he grows toward reading maturity." The program is built upon a framework of experiences resulting in pupil- and teacher-made materials. The concept underlying the program is that children's language development proceeds from oral expression through written expression thus creating high motivation for reading one's own materials and easy transfer to reading what others have written. Allen states: "Utilization of the child's language as a basis of reading instruction results in a high degree of independence in writing and reading" (1964, p. 63).

The language experience approach rejects the idea of a controlled vocabulary for beginning readers, and the development of a basic sight vocabulary is considered an individual matter based upon the child's oral expression. "The direct teaching program for phonics and other word recognition skills is more closely related to the writing and spelling activities where children are dealing with the language letter by letter, syllable by syllable, and word by word" (Allen, 1964, p. 64). It is claimed that dictation and writing their own stories enables children to recognize enough words that they can read material written by others. As the children develop their skill in reading, they select their own reading material.

There is evidence concerning the interrelationship of all the communication skills (speaking, listening, writing, and reading) and this is acknowledged by most reading authorities regardless of the approach they advocate for beginning reading instruction. Gray summarized the situation as follows:

> Summaries of research by Hildreth and by other specialists have shown that reading and the other language arts are closely interrelated in many important respects. It has been proposed, therefore, that instruction in all the language arts should be provided in a closely integrated program. Although many efforts have been made to develop such a program, no carefully controlled studies of its advantages and limitations have been reported (1960, p. 1117).

Loban (1963) concluded from his longitudinal study of children's language that the children who were high in general language ability, based on teachers' ratings of oral language and vocabulary scores, were also high in reading ability. The children who were low in general language ability were also low in reading ability.

In a five year comparative study of the basic approach, individualized approach, and the language experience approach, R.V. Allen (1962) found that children taught by the language experience approach made as much or more progress in reading, as measured on standardized tests, than did pupils taught through individualized and basic approaches.

There is little doubt that reading and other language skills are related. However, the research evidence concerning the value of language experience methods is sparse and more research is needed before any conclusions can be drawn.

Sex differences and reading

Research evidence concerning sex differences in reading achievement generally favors the girls. There are numerous theories as to the cause of these results, but, to date, there has been no conclusive evidence as to the causes of these differences.

Balow (1963), in a study of 151 girls and 151 boys with equivalent mean IQs found that the girls were superior to the boys in a reading readiness test. However, when reading readiness was held constant, using an analysis of covariance, no significant differences were found between the reading achievement of boys and girls at the end of Grade 1. Balow inferred from this study that the data supported the non maturational, cultural theory of sex differences in reading achievement because perception and readiness appear to be affected by training.

In a study of sex differences in reading readiness at the first-grade level, Carroll (1948) found that girls were slightly superior to boys in tests of visual, auditory, language and articulation ability, and of ability to name letters.

Gates (1961b) studied sex differences in reading ability of 13,114 subjects (6,646 boys and 6,648 girls) in Grades 2 through 8 in 12 school systems in 10 states. On 21 comparisons made on tests of speed, vocabulary, and level of comprehension, the mean raw scores for girls were higher than those for boys. Gates concluded that, on the average, girls' reading abilities exceeded those of boys. He commented that maturity did not explain the superiority of the girls because, in his study, the girls were superior in the upper grades as well as in the lower ones. Environmental, rather than hereditary, factors were suggested as causes for the differences in achievement.

As a group, the boys among 1,500 second-grade pupils studied by Pauley (1951) were 2 months older chronologically than the girls, but their mean reading achievement was 2 months below that of the girls.

Templin (1957) reported that girls were superior in articulation and sound discrimination at the age of 8, while boys were superior in vocabulary at the ages of 6 to 8. Few significant differences were apparent between boys and girls at the ages of 3 through 5 years.

More boys than girls become remedial reading cases, as shown by Heilman (1961) who reported the data shown in Table 1 from a number of studies showing the percentage of boys and girls referred as remedial reading cases.

In a study of reading achievement of German and American children, Preston (1962) matched 1,338 children in Philadelphia with 1,053 children in Wiesbaden, Germany. The children were matched on intelligence, parental occupations, and instructional level. The children were tested using cross-translations on the Gates Reading Survey and the Frankfurter Test. Preston reported the German children were generally lower in comprehension than the American children. However, the

Table 1 Sex differences in remedial reading cases (Heilman, 1961, p. 365)

Study	Date	No. of cases		Percent	
		Boys	Girls	Boys	Girls
Blanchard	1936	63	10	86	14
Young	1938	37	4	90	10
Preston	1940	72	28	72	28
Missildine	1946	25	5	83	17
McCollum & Shapiro	1947	31	9	76	24
Axline	1947	28	9	76	24
Vorhaus	1952	178	47	80	20
Johnson	1955	23	11	67	33
Fry	1959	163	39	81	19

difference was less at the sixth grade than at the fourth grade, and there was no difference for the sixth-grade boys. German boys were superior to German girls in reading ability, adding support to the theory that environmental conditions are causing the sex differences favoring girls in America.

Waetjen and Grambs (1963) have suggested that schools reward verbal comprehension and language skill, consequently reinforcing girls' greater facility with language. As a result of receiving little reward, the boys feel negative about their adequacy with language skills. Thus, language activities become identified as girl-like activities with the result that boys cannot then participate as fully as they might have in activities involving language.

It is a fairly well established conclusion that girls are superior to boys in reading achievement as well as general language ability. There is some doubt, however, as to the causes for sex differences in reading. In this regard, there is some evidence to support the theory that the causes of sex differences in reading are related to environmental conditions within the society and the schools.

Summary of review literature

It is evident, from the perusal of the studies reviewed that little conclusive evidence has been reported concerning the comparative efficiency of the methods of teaching reading. Some of the methods with which this report is concerned are new approaches and have not been thoroughly researched. In reviewing the studies cited here, the experimental methods are often variations of a general class of methods and not totally equivalent, limiting the comparability of the conclusions reported. There is also the limitation of the effect of uncontrolled variables which may have confounded the results reported in some of the studies.

In conclusion, the superiority of a single method of reading instruction is yet to be determined. It appears that a composite of methods would produce the best results and that an effort should be made to determine what each method would contribute to the reading program.

Overview of Cooperative Research Program in First-Grade Reading Instruction

Each of the 27 studies comprising the Cooperative Research Program in First-Grade Reading Instruction was a complete study in itself. A brief description of each of the 27 projects is presented in Appendix I. Each was selected on the basis of its potential for yielding valuable information about the teaching of beginning reading. The program's unique characteristic was that each project director, in addition to carrying out his own analysis, made the data available to the Coordinating Center, thus enabling an analysis across individual projects. Most of the projects investigated instructional methodology and the evaluation of method is the major focus of the present report. A number of projects concerned themselves with aspects of the instructional program in beginning reading other than methodology.

Some of the Cooperative Research Program studies not concerned with instructional materials investigated various grouping plans. One project (Sister M. Marita) evaluated the relative effectiveness of a "whole-class" system in which all pupils in the room met as a single group. This approach had the proposed advantage of increasing each child's contact time with the teacher since it was not necessary for the teacher to divide time among three groups. Another researcher (Wyatt) investigated the effectiveness of grouping beginning readers by sex on the assumption that girls constitute unfair competition and tend to dominate the typical heterosexual reading group.

Other projects in the study investigated various devices for helping the beginning teacher of reading. One (Heilman) studied the effect of an intensive inservice program on teachers' classroom behavior and reading achievement of pupils taught by the experimental teachers. The inservice program consisted of a 2-week preschool seminar and 25 two-hour seminar sessions held during the first 30 weeks of the school year. Another study (Morrill) sought to determine the feasibility of improving the reading achievement of first-grade children by utilizing consultants in two different ways: one used the typical technique of consultant help on a one-to-one basis in which the consultant answers a request for his services from the teacher or building principal; the other used a method by which the consultant brought together teachers with common problems such as those found in first-grade reading instruction in scheduled meetings on released time. The second approach was designed to foster interaction among teachers.

A brief description of each of the 27 projects composing the Cooperative Research Program in First-Grade Reading Instruction is presented in Appendix I.

Procedures

To enable comparison between the 27 studies conducted under the Cooperative Research Program in First-Grade Reading Instruction, common procedures for data collection and analysis as well as common experimental procedures were established. The role of the Coordinating Center was crucial in this regard.

Role of the Coordinating Center

The Coordinating Center was established to perform two functions. First, the Center was charged with the responsibility for maintaining communication among the various projects and for facilitating thereby the cooperative aspects of the study. Its first function, therefore, was to host a conference of the individual project directors in June, 1964, at which decisions were made concerning experimental procedures and data collection. At this meeting, the directors decided upon common prereading and reading outcome measures to be used by all projects. They also agreed to collect information common to all studies about teacher, pupil, school, and community characteristics which might reasonably be expected to be related to success or failure in beginning reading.

Two further meetings of project directors were held—the first in December, 1964, at the University of Minnesota; the second during the International Reading Association Convention in Detroit during May, 1965. These meetings were devoted to discussions of problems concerning cooperative aspects of the study. Every effort was made to establish experimental controls common to all projects in order to make possible comparisons between and among individual studies.

Uniformity in procedures was further enhanced through periodic memoranda issued by the Coordinating Center. A common format for recording data on cards was devised to facilitate the analysis of the common data. The Center also served as a clearinghouse for questions about administration or scoring of certain of the tests employed in the study. In addition, all but five of the projects were visited by either the director or associate director of the Center. These visits enabled Center staff to get a first hand look at each project in order to be in a better position to interpret data collected from them. Moreover, the visits provided an opportunity for the individual project director to discuss any problems he might have relative to the cooperative aspects of the research.

The second major function of the Coordinating Center was to collect, organize, analyze, and interpret the data common to each child in all 27 individual projects. This function, of course, is the basis for the present report.

Data collected

A great deal of information about each pupil who participated in each of the 27 studies, about his teacher, about the class and school in which he was enrolled, and about the community in which the pupil lived was collected by all of the participating project directors.

Pupil data. For each pupil, information was gathered concerning sex, chronological age at the beginning of the year, amount of preschool experience, and the number of days absent during the experimental period.

Data regarding the child's readiness for reading were gathered by means of an intelligence test and various tests of auditory discrimination, visual discrimination, and language facility. The group intelligence test was the Pintner-Cunningham Primary Test. Reading readiness information was gathered by administering: (1) the Murphy-Durrell Phonemes Test, which tests the ability to discriminate between like and unlike sounds; (2) the Murphy-Durrell Letter Names Test, which tests the child's ability to recognize lower case and capital letters; (3) the Murphy-Durrell Learning Rate Test, which tests the child's ability to learn a small number of words; (4) the Thurstone Pattern Copying Test, which tests the child's ability to copy a figure; (5) the Thurstone-Jeffrey Identical Forms Test, which asks the child to select from a group of figures a figure similar to one used as a stimulus; (6) the Metropolitan Word Meaning Test, which is essentially a vocabulary test; (7) the Metropolitan Listening Test, which measures a child's ability to follow directions. In addition to these tests which were given to all pupils, the Detroit Word Recognition Test was administered to those pupils who gave some evidence of being able to read at the beginning of first grade.

Postinstructional tests were selected to measure silent and oral reading ability as well as spelling ability, writing ability, and attitude toward reading. The group-administered Stanford Achievement Test, Primary Battery I, was administered to all students. Five subtests were used to measure the child's reading and general language ability. These subtests were: (1) the Word Reading Test, consisting of 35 items, which measures the ability of a pupil to identify a word without the aid of context; (2) the Paragraph Meaning Test, which is a measure of the child's ability to comprehend connected discourse ranging in length from single sentences to paragraphs of six sentences, and which involves levels of comprehension varying from extremely simple recognition to the making of inference from several related sentences; (3) the Vocabulary Test, which measures pupils' vocabulary independent of reading skill; (4) the Spelling Test, which is a dictation-type exercise; and (5) the Word Study Skills Test, which tests auditory perception and phonics ability.

In addition to the group-administered Stanford test of silent reading ability, a sample from each treatment group within each project was administered the Gilmore Oral Reading Test. This sample consisted of 20 to 50 students randomly selected from each treatment group. The Gilmore test was scored in terms of reading accuracy and reading rate. The same sample pupils were asked to pronounce words from the Gates Word Pronunciation Test and the Fry Phonetically Regular Words Test. The Gates test consisted of the first two columns from the Gates Diagnostic Reading Test. These words are listed according to increasing difficulty, but there is no attempt to control sound-symbol regularity in the gradation of the words. The Fry test is a list of words controlled on the basis of sound-symbol relationships and graded roughly in order of difficulty by vowel sounds used—short vowel words, long vowel words, broad *a*, vowel modified by *r*, and the like. In each of these word lists, the child reads aloud and pronounces the word without the benefit of context.

Measures of the child's writing ability and his attitude toward reading were also obtained. The San Diego Pupil Attitude Inventory was administered to all pupils in those projects which chose to employ this instrument. Because not every project director administered this test, the analysis of the combined data reported in the current study does not include this attitude measure. The sample pupils who were administered the Fry and Gates word lists and the Gilmore Oral Reading Test, were also asked to write a story from a stimulus common to all projects. The writing sample was to be evaluated in terms of mechanics and creative expression. However, because of the difficulty of scoring, not all projects made use of this evaluative technique; and, therefore, the analysis of the combined data does not include this variable. However, reports of the various individual projects may include pertinent information concerning writing ability as it is related to different instructional programs.

In summary, the pupil information that was utilized in the analysis of combined data reported in the current study included the seven readiness measures, the intelligence test, the five Stanford Achievement subtests, the two Gilmore Oral Reading measures, and the Fry and Gates word lists. Information concerning sex was also utilized in that all of the analyses were run using sex as a blocking variable. Pupil data not included in the analysis of the combined data included chronological age, which proved to be unrelated to reading achievement; amount of preschool experience, which was categorized in such a fashion as to make it impossible to use in a covariance analysis; number of days absent during the experimental period, which was found to be unrelated to reading achievement; and the attitude and writing measures, which were excluded for reasons already given.

Teacher data. Data were collected concerning each teacher's (1) sex, (2) age, (3) degrees earned, (4) certification, (5) years of teaching experience, (6) years of experience teaching first grade, (7) marital status, (8) number of children, (9) attitude toward teaching of reading as measured by the San Diego Teacher Attitude Scale, (10) number of days absent during the experimental period, and (11) teaching effectiveness as rated by supervisors. All of these data are reported elsewhere (Bond & Dykstra, 1967, Appendix C), although only years of experience was utilized as a covariate in any of the analyses. Quantitative measures, such as number of days absent and score on the San Diego Teacher Attitude Scale, proved to be unrelated to the reading achievement of pupils and, therefore, were not used. The categorical data which were analyzed, such as the type of teaching certificate held, likewise proved to be relatively unrelated to pupil achievement in reading. The teacher efficiency rating was not utilized because of lack of objectivity which raised questions about reliability and validity, and because it was related to only a slight degree to pupil success in reading.

School and community data. Information collected about community characteristics included median education of adults in the community, median income of adults according to census figures, population of the community, and type of community (urban, rural, or suburban). Information collected about schools included the number enrolled in each first-grade class, length of the school day, length of the school year, number of first-grade rooms in the building, number of first-grade rooms in the district, whether or not the school had the services of a school librarian, and the per pupil costs for education. These data for each project are also included in the Appendix C to the Coordinating Center's final report (Bond & Dykstra, 1967). No further reference is made to them in terms of the analysis. In the first place, there was little indication that any of the school and community characteristics were significantly related to the reading achievement. This statement of no relationship, of course, is valid only with reference to the specific communities, schools, and school populations included in this project. Furthermore, many of the community characteristics were categorized in such a way as to make it difficult to use them as control variables in a covariance analysis.

Common experimental guidelines

In addition to administering common preinstructional and postinstructional tests and collecting common information about teachers, schools, and communities, the project directors also agreed to abide by certain experimental guidelines. These were necessary, of course, to make possible comparisons between studies. The following procedural controls were considered essential: (1) All testing instruments to be used in the collection of the data were not to be in the hands of the classroom teacher until the close of the school day preceding the day the test was to be given. (2) Tests were not to be scored by the classroom teacher although she could administer the tests if the building principal or other professional person acted as an observer. (3) No instructions were to be given to the classroom teacher in test procedures beyond those which were provided in the manual for a given test. (4) The length of the experimental program was designated to be 140 instructional days. Pretests and posttests were to be given before and after this 140-day period. Final testing was to begin on the 141st day regardless of the time of year. (5) Each project director was encouraged to take whatever steps would be necessary to control for the Hawthorne Effect which would probably be associated with novel experimental programs.

Organization of the data

The Coordinating Center devised a format to be used by all project directors in recording common data collected. Each project director then punched two sets of data cards, one for his own use and one for the Coordinating Center. Three cards were punched for each pupil in the study. The first of these cards included data on the readiness characteristics of the pupils and the data regarding teacher, school, and community characteristics which could be obtained at the beginning of the school year. The second card for each child recorded the data on the outcome measures and data such as class size at the end of the year which could be obtained at the end of the experimental period. The third card recorded unique data which the project director had collected. Only the first two cards, those which carried common data, were organized and analyzed by the Center.

When the data cards arrived at the Coordinating Center, they were first run through a computer program which eliminated from the analysis all pupils on whom complete data were not available. Any pupil's scores were included in the analysis only if he had taken all seven of the readiness tests, the Pintner-Cunningham Primary Test, and all five Stanford Achievement tests. Furthermore, data were used in the analysis only if information concerning the child's sex and chronological age were punched on his card. As a result, varying numbers of pupils in the various projects were eliminated from the combined analysis because of missing data. However, the pupils eliminated for this reason were relatively few in number. The pupils with complete data were used in the analyses discussed later.

After the cards were screened to eliminate pupils with missing data, descriptive statistics were calculated for all of the quantitative and most of the categorical data. For these descriptive statistics, individuals were used as the experimental unit. The statistics were calculated separately for boys and girls within treatment and within project. These descriptive data are tabled in Appendix A to the Coordinating Center's final report (Bond & Dykstra, 1967). The tables reveal the differences among projects with respect to pupil, teacher, school, and community characteristics. They also point out differences within projects between treatments on the same variables. The descriptive data illustrate very graphically the tremendous project differences in reading achievement of pupils, in prereading readiness characteristics of pupils, and in various teacher, school, and community characteristics. They also show that many times the projects were unsuccessful in assigning pupils of equal ability to each of the various treatment groups.

General procedure of analysis

The present investigation was designed to obtain information relevant to three basic questions: (1) To what extent are various pupil, teacher, class, school, and community characteristics related to pupil achievement in first-grade reading and spelling? (2) Which of the many approaches to initial reading instruction produces superior reading and spelling achievement at the end of the first grade? (3) Is any program uniquely effective or ineffective for pupils with high or low readiness for reading?

In order to assess the relationships between various pupil, teacher, class, school, and community characteristics and subsequent pupil achievement in reading, product-moment correlation coefficients were computed. Information about the numbers of pupils involved and the results of this analysis are reported later.

Comparisons of method are discussed in succeeding sections as are the major techniques of analysis used in the present report. In this section of the analysis, various reading programs were evaluated by comparing their effectiveness with that of well-known basal readers used in the same project. Extensive project by treatment interactions, extensive project effects for treatment, and lack of complete replication of treatments in all projects made this the most appropriate technique of analysis to use. Procedures are discussed more completely later along with a presentation of the results.

The third general purpose for the study was to determine whether or not any of the programs was uniquely effective for pupils with high or low readiness for reading. Readiness for reading in this portion of the analysis was assessed by means of an intelligence test, a measure of auditory discrimination, and a test of letter knowledge. Pupils were blocked in turn according to their performance on each of the three measures. Then the appropriate treatment by readiness characteristic interaction was examined to note whether or not it could have occurred by chance. A significant interaction would indicate that treatments were not operating in the same manner across all ranges of readiness. Discussion of the procedures and results for this analysis are presented later.

Analysis of relationships

A discussion of the relationships between reading and spelling achievement at the end of the first grade and (1) pupil characteristics such as chronological age, mental age, number of days absent, and readiness for reading; (2) teacher characteristics such as years of teaching experience, years of experience teaching first grade, efficiency rating, and days absent; and (3) class size follows. Relationships among the various individual outcome measures and the group-administered Stanford Achievement Test were assessed. The Pearson product-moment correlation coefficient was utilized in all cases.

Relationships between readiness and reading

Complete data were gathered on seven reading readiness measures and the Pintner-Cunningham Intelligence Test. Achievement was measured by the five subtests of the Stanford Achievement Test. Product-moment correlation coefficients between each prereading measure and each achievement measure were computed separately for each of the treatments identified as Basal, Basal plus Phonics, i.t.a., Linguistic, Phonic/Linguistic, and Language Experience. Each of the correlations was calculated by pooling within class and sex for relevant projects. The number of pupils on whom the correlations were based varied from treatment to treatment.

Correlation relationships for Basal treatment. The product-moment correlation coefficients among readiness measures, among outcome measures, and between each readiness measure and each outcome measure for the Basal treatment are reported in Table 2. In general, the intercorrelations of the prereading measures range from .20 to .40, thereby indicating that these tests appear to measure different facets of readiness. Relationships among the achievement measures, on the other hand, are somewhat higher with the correlation coefficient between word recognition and paragraph meaning .76.

The best single predictor of achievement on the Stanford Achievement battery was the Murphy-Durrell Letter Names Test. This test correlated .55 with Word Reading, .52 with Paragraph Meaning, .41 with Vocabulary, .48 with Spelling, and .51 with Word Study Skills. The Murphy-Durrell Phonemes Test also correlated substantially with the achievement measures. The other reading readiness subtests correlated .40 or less with the reading and spelling measures. The Pintner-Cunningham Primary Intelligence Test was related to the reading achievement measures to a somewhat lesser extent than the Phonemes and Letter Names Tests. For example, the correlation between the intelligence test and the Paragraph Meaning subtest was .42.

Correlation relationships for the i.t.a. treatment. Intercorrelations for the same variables are reported for the i.t.a. treatment in Table 3. Again, the Letter Names subtest was the best predictor of future success on the Stanford Achievement Test. The Letter Names Test correlated .60 with Word Reading, .58 with Paragraph Meaning, .48 with Vocabulary, .53 with Spelling, and .59 with Word Study Skills. The Phonemes subtest and the Pintner-Cunningham Intelligence Test also correlated to a relatively high degree with the Stanford measures. The correlations between prereading measures and reading achievement measures were found to be somewhat higher for the i.t.a. group than for the Basal group but, in general, were very similar.

Table 2 Correlation matrix for the Basal treatment[a]

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Murphy-Durrell phonemes		.50	.37	.35	.29	.35	.26	.46	.48	.46	.44	.37	.48
2 Murphy-Durrell letter names	.50		.45	.31	.30	.34	.21	.43	.55	.52	.41	.48	.51
3 Murphy-Durrell learning rate	.37	.45		.28	.27	.26	.25	.34	.38	.40	.28	.32	.35
4 Thurstone pattern copying	.35	.31	.28		.32	.26	.25	.49	.34	.34	.32	.30	.36
5 Thurstone-Jeffrey identical forms	.29	.30	.27	.32		.28	.24	.46	.29	.29	.32	.26	.31
6 Metropolitan word meaning	.35	.34	.26	.26	.28		.46	.40	.32	.30	.45	.25	.32
7 Metropolitan listening	.26	.21	.25	.25	.24	.46		.34	.22	.23	.31	.16	.25
8 Pintner-Cunningham raw score	.46	.43	.34	.49	.46	.40	.34		.44	.42	.50	.32	.44
9 Stanford word reading	.48	.55	.38	.34	.29	.32	.22	.44		.76	.51	.63	.70
10 Stanford paragraph meaning	.46	.52	.40	.34	.29	.30	.23	.42	.76		.49	.66	.69
11 Stanford vocabulary	.44	.41	.28	.32	.32	.45	.31	.50	.51	.49		.40	.56
12 Stanford spelling	.37	.48	.32	.30	.26	.25	.16	.32	.63	.66	.40		.64
13 Stanford word study skills	.48	.51	.35	.36	.31	.32	.25	.44	.70	.69	.56	.64	
Means	25.2	32.7	9.4	13.6	15.2	8.5	9.0	36.6	19.6	19.7	21.1	11.1	35.5
Standard deviations	11.29	12.05	3.95	6.08	6.48	2.70	2.90	7.86	6.03	7.61	5.72	5.36	8.71

[a] Correlations were calculated by pooling within class and sex for those treatments labeled Basal. Means and pooled estimates of the standard deviations are presented at the base of the table. The N upon which this table is based is 4,266 from 187 classes in 17 projects.

Table 3 Correlation matrix for the i.t.a. treatment[a]

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Murphy-Durrell phonemes		57	36	42	25	39	31	50	54	53	52	43	54
2. Murphy-Durrell letter names	57		37	38	28	37	30	48	60	58	48	53	59
3. Murphy-Durrell learning rate	36	37		26	20	23	22	31	35	34	26	30	30
4. Thurstone pattern copying	42	38	26		27	33	30	51	45	41	40	35	43
5. Thurstone-Jeffrey identical forms	25	28	20	27		30	27	40	36	33	34	25	33
6. Metropolitan word meaning	39	37	23	33	30		38	48	40	38	51	28	41
7. Metropolitan listening	31	30	22	30	27	38		38	31	29	41	25	31
8. Pintner-Cunningham raw score	50	48	31	51	40	48	38		52	52	58	40	52
9. Stanford word reading	54	60	35	45	36	40	31	52		83	60	64	77
10. Stanford paragraph meaning	53	58	34	41	33	38	29	52	83		60	60	75
11. Stanford vocabulary	52	48	26	40	34	51	41	58	60	60		42	62
12. Stanford spelling	43	53	30	35	[illegible]	28	25	40	64	60	42		66
13. Stanford word study skills	54	59	30	43	33	41	31	52	77	75	62	66	
Means	28.9	31.7	9.9	12.0	15.7	9.1	9.2	39.1	23.5	21.4	22.0	10.6	38.8
Standard deviations	11.49	12.32	4.41	6.50	6.45	2.52	2.46	7.71	6.98	9.54	5.89	5.13	8.96

[a] Correlations were calculated by pooling within class and sex for those treatments labeled i.t.a. Means and pooled estimates of the standard deviations are presented at the base of the table. The N upon which this table is based is 1,055 from 48 classes in five projects.

Table 4 Correlation matrix for the Basal plus Phonics treatment[a]

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Murphy-Durrell phonemes		59	45	44	41	46	41	56	54	52	47	46	55
2. Murphy-Durrell letter names	59		50	47	46	46	43	58	58	55	46	53	56
3. Murphy-Durrell learning rate	45	50		35	35	32	33	45	39	39	32	34	37
4. Thurstone pattern copying	44	47	35		45	34	37	59	46	46	38	40	47
5. Thurstone-Jeffrey identical forms	41	46	35	45		34	35	54	40	40	39	37	41
6. Metropolitan word meaning	46	46	32	34	34		49	54	41	44	52	34	41
7. Metropolitan listening	41	43	33	37	35	49		51	37	38	42	31	39
8. Pintner-Cunningham raw score	56	58	45	59	54	54	51		57	56	54	51	57
9. Stanford word reading	54	58	39	46	40	41	37	57		77	55	70	76
10. Stanford paragraph meaning	52	55	39	46	40	44	38	56	77		55	71	73
11. Stanford vocabulary	47	46	32	38	39	52	42	54	55	55		48	58
12. Stanford spelling	46	53	34	40	37	34	31	51	70	71	48		73
13. Stanford word study skills	55	56	37	47	41	41	39	57	76	73	58	73	
Means	21.4	30.8	8.1	12.3	12.0	7.5	7.9	34.1	21.1	20.4	21.5	10.8	35.9
Standard deviations	12.65	14.14	4.69	6.72	7.07	2.76	2.75	9.09	6.49	8.54	5.58	5.54	9.32

[a] Correlations were calculated by pooling within class and sex for those treatments labeled Basal plus Phonics. Means and pooled estimates of the standard deviations are presented at the base of the table. The N upon which this table is based is 1,104 from 40 classes in four projects.

Correlation relationships for the Basal plus Phonics treatment. Intercorrelations for the Basal plus Phonics treatment are presented in Table 4. One of the best predictors of achievement on the Stanford was again the Letter Names Test. The Letter Names Test correlated .58 with Word Reading, .55 with Paragraph Meaning, .46 with Vocabulary, .53 with Spelling, and .56 with Word Study Skills. Correlations between the Pintner-Cunningham test and the Stanford Achievement Test were of approximately the same magnitude. The Phonemes test also correlated well with the criterion measures. All of the predictive validity coefficients are somewhat higher for this treatment than for the Basal treatment. However, the tests tend to rank in approximately the same order as far as their predictive validity is concerned.

Correlation relationships for the Language Experience treatment. The correlations between prereading measures and reading achievement measures for the Language Experience approach are found in Table 5. As a group, these correlations are somewhat lower than the correlations found for previous treatments. Again, the Letter Names Test was the best single predictor of future success in reading and spelling. Knowledge of letter names correlated .52 with Word Reading, .51 with Paragraph Meaning, .36 with Vocabulary, .53 with Spelling, and .48 with Word Study Skills. These correlation

Table 5 Correlation matrix for the Language Experience treatment[a]

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Murphy-Durrell phonemes		48	35	34	24	30	27	42	45	41	40	38	44
2 Murphy-Durrell letter names	48		35	35	26	28	24	40	52	51	36	53	48
3 Murphy-Durrell learning rate	35	35		22	22	34	33	26	29	28	19	30	26
4 Thurstone pattern copying	34	35	22		30	16	19	51	33	33	29	34	37
5 Thurstone-Jeffrey identical forms	24	26	22	30		21	23	41	29	29	30	27	30
6 Metropolitan word meaning	30	28	34	16	21		58	28	21	19	33	16	21
7 Metropolitan listening	27	24	33	19	23	58		26	17	18	25	16	21
8 Pintner-Cunningham raw score	42	40	26	51	41	28	26		42	43	43	38	45
9 Stanford word reading	45	52	29	33	29	21	17	42		71	44	71	73
10 Stanford paragraph meaning	41	51	28	33	29	19	18	43	71		50	70	69
11 Stanford vocabulary	40	36	19	29	30	33	25	43	44	50		42	50
12 Stanford spelling	38	53	30	34	27	16	16	38	71	70	42		73
13 Stanford word study skills	44	48	26	37	30	21	21	45	73	69	50	73	
Means	24.7	34.3	9.8	16.0	15.7	9.0	9.3	38.4	21.0	20.0	22.1	11.8	36.4
Standard deviations	11.88	11.36	4.81	6.36	7.11	2.89	3.03	7.50	6.60	8.53	5.86	4.85	8.49

[a] Correlations were calculated by pooling within class and sex for those treatments labeled Language Experience. Means and pooled estimates of the standard deviations are presented at the base of the table. The N upon which this table is based is 1,431 from 60 classes in four projects.

Table 6 Correlation for the Linguistic treatment[a]

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Murphy-Durrell phonemes		47	33	35	26	38	29	46	55	50	48	51	57
2 Murphy-Durrell letter names	47		42	25	17	30	23	35	56	55	40	56	51
3 Murphy-Durrell learning rate	33	42		26	18	19	16	34	44	48	28	40	39
4 Thurstone pattern copying	35	25	26		28	21	17	49	36	35	27	33	35
5 Thurstone-Jeffrey identical forms	26	17	18	28		19	17	44	29	27	24	26	26
6 Metropolitan word meaning	38	30	19	21	19		33	36	31	27	43	26	31
7 Metropolitan listening	29	23	16	17	17	33		31	23	27	32	18	26
8 Pintner-Cunningham raw score	46	35	34	49	44	36	31		47	48	47	41	46
9 Stanford word reading	55	56	44	36	29	31	23	47		75	52	78	74
10 Stanford paragraph meaning	50	55	48	35	27	27	27	48	75		46	72	66
11 Stanford vocabulary	48	40	28	27	24	43	32	47	52	46		44	53
12 Stanford spelling	51	56	40	33	26	26	18	41	78	72	44		75
13 Stanford word study skills	57	51	39	35	26	31	26	46	74	66	53	75	
Means	21.7	20.7	9.4	10.3	15.3	8.3	8.5	33.7	18.7	16.2	20.0	9.4	34.3
Standard deviations	10.94	10.83	3.63	5.69	7.85	2.53	2.29	8.17	6.29	7.33	5.15	4.79	7.76

[a] Correlations were calculated by pooling within class and sex for those treatments labeled Linguistic. Means and pooled estimates of the standard deviations are presented at the base of the table. The N upon which this table is based is 700 from 31 classes in three projects.

coefficients are not substantially different from those obtained between similar variables for the other treatments.

Correlation relationships for the Linguistic treatment. The intercorrelations for the Linguistic treatment are presented in Table 6. The Letter Names and Phonemes subtests were the best predictors of achievement. In general, the correlation coefficients looked very much like those reported for the other treatments.

Correlation relationships for the Phonic/Linguistic treatment. The intercorrelations for the Phonic/Linguistic treatment are reported in Table 7. The three best predictors of success again were the Letter Names, Phonemes, and Pintner-Cunningham tests. The Letter Names subtest correlated .57 with Word Reading, .59 with Paragraph Meaning, .47 with Vocabulary, .54 with Spelling, and .55 with Word Study Skills. The Phonemes and Pintner-Cunningham tests also correlated near or above .50 with the criterion measures.

Summary of relationships between readiness and reading. A summary of predictive relationships of the various prereading measures is reported in Tables 8 and 9. In Table 8, the Paragraph Meaning subtest of the Stanford Achievement Test is used as a measure of reading achievement. For each of the treatment groups, the Murphy-Durrell Letter Names Test ranked first in its relationship with the criterion. Likewise, in four of the six

Table 7 Correlation matrix for the Phonic/Linguistic treatment[a]

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Murphy-Durrell phonemes		.56	.54	.40	.33	.41	.40	.54	.59	.57	.53	.50	.59
2. Murphy-Durrell letter names	.56		.49	.27	.30	.34	.33	.50	.57	.59	.47	.54	.55
3. Murphy-Durrell learning rate	.54	.49		.35	.28	.30	.27	.44	.48	.52	.40	.45	.50
4. Thurstone pattern copying	.40	.27	.35		.28	.27	.20	.50	.36	.34	.34	.35	.39
5. Thurstone-Jeffrey identical forms	.33	.30	.28	.28		.35	.32	.52	.38	.35	.31	.32	.35
6. Metropolitan word meaning	.41	.34	.30	.27	.35		.39	.48	.35	.32	.55	.29	.36
7. Metropolitan listening	.40	.33	.27	.20	.32	.39		.51	.31	.33	.42	.27	.31
8. Pintner-Cunningham raw score	.54	.50	.44	.50	.52	.48	.51		.56	.52	.54	.47	.56
9. Stanford word reading	.59	.57	.48	.36	.38	.35	.31	.56		.81	.60	.84	.80
10. Stanford paragraph meaning	.57	.59	.52	.34	.35	.32	.33	.52	.81		.60	.76	.76
11. Stanford vocabulary	.53	.47	.40	.34	.31	.55	.42	.54	.60	.60		.60	.63
12. Stanford spelling	.50	.54	.45	.35	.32	.29	.27	.47	.84	.76	.60		.80
13. Stanford word study skills	.59	.55	.50	.39	.35	.36	.31	.56	.80	.76	.63	.80	
Means	29.2	35.6	10.8	9.6	15.3	9.1	9.5	39.5	26.7	24.4	24.1	14.2	41.7
Standard deviations	12.12	11.55	3.91	5.77	5.55	2.53	2.29	7.30	6.44	9.33	6.08	5.49	8.98

[a] Correlations were calculated by pooling within class and sex for those treatments labeled Phonic/Linguistic. Means and pooled estimates of the standard deviations are presented at the base of the table. The N upon which this table is based is 488 from 23 classes in three projects.

Table 8 Summary of correlations between premeasures and Stanford Paragraph Meaning Test for each of six treatments

	Basal	i.t.a	Basal plus Phonics	Language Experience	Linguistic	Phonic/Linguistic
Murphy-Durrell phonemes	.46	.53	.52	.41	.50	.57
Murphy-Durrell total letters	.52	.58	.55	.51	.55	.59
Murphy-Durrell learning rate	.40	.34	.39	.28	.48	.52
Thurstone pattern copying	.34	.41	.46	.33	.35	.34
Thurstone-Jeffrey identical forms	.29	.33	.40	.29	.27	.35
Metropolitan word meaning	.30	.38	.44	.19	.27	.32
Metropolitan listening	.23	.29	.38	.18	.27	.33
Pintner-Cunningham Intelligence Test	.42	.52	.56	.43	.48	.52

treatment groups, the Murphy-Durrell Phonemes Test ranked as the second best predictor of reading achievement. The lowest correlation between Letter Names and the Paragraph Meaning subtest was .51, while the highest relationship was .59. Obviously, the ability to recognize letters at the beginning of first grade was related to reading success in all of the methods and programs employed in the study.

Correlations between the other readiness measures and reading achievement were more variable. Coefficients of correlation between the Phonemes subtest and the Paragraph Meaning subtest ranged from .41 to .57. Furthermore, correlations with Paragraph Meaning ranged from .28 to .52 for the Learning Rate test, .33 to .46 for the Pattern Copying test, .27 to .40 for the Identical Forms test, .19 to .44 for the Metropolitan Meaning test, .18 to .38 for the Metropolitan Listening test, and .42 to .56 for the Pintner-Cunningham Intelligence Test. For these tests, there was some indication that predictive relationships were higher within some treatments than within others.

The predictive relationship of each of the various subtests with the Word Reading tests are presented in Table 9. For five of the six treatments, the Letter Names subtest was the best predictor of Word Recognition ability. The only exception was the Phonic/Linguistic treatment where the Phonemes test correlated best with the criterion. The lowest correlation between Letter Names and Word Reading was .52 for the Language Experience group, while the highest correlation was .60 for the i.t.a. group. For most of the other readiness measures, the predictive relationship was also consistent from treatment to treatment.

Table 9 Summary of correlations between premeasures and Word Reading Test for each of six treatments

	Basal	i.t.a	Basal plus Phonics	Language Experience	Linguistic	Phonic/Linguistic
Murphy-Durrell phonemes	.48	.54	.54	.45	.55	.59
Murphy-Durrell letter names	.55	.60	.58	.52	.56	.57
Murphy-Durrell learning rate	.38	.35	.39	.29	.44	.48
Thurstone pattern copying	.34	.45	.46	.33	.36	.36
Thurstone-Jeffrey identical forms	.29	.36	.40	.29	.29	.38
Metropolitan word meaning	.32	.40	.41	.21	.31	.35
Metropolitan listening	.22	.31	.37	.17	.23	.31
Pintner-Cunningham Intelligence Test	.44	.52	.57	.42	.47	.56

Intercorrelations among group and individual achievement measures

Intercorrelations were also computed for the Stanford Word Reading, Stanford Paragraph Meaning, Gilmore Accuracy, Gilmore Rate of Reading, Fry Phonetically Regular Word List, and Gates Word Pronunciation Test. There was special interest in determining the relationships among the Stanford Word Reading Test, the Fry Word List, and the Gates Word List. The Stanford test differed from the other two in that it was administered to a group. The Fry test was unique in that it was designed to test children's recognition of phonetically regular words. The Gates list, on the other hand, consisted of high frequency words with no regard for regularity in sound-symbol relationships. The investigators were interested in assessing the degree of independence among these three measures of word recognition. Correlation coefficients computed for each treatment group are reported in Tables 10, 11, 12, 13, 14, and 15. The correlations between the Stanford Word Reading Test and the Fry test ranged from .69 to .83 for the six treatments. Correlations between the Word Reading test and the Gates test varied from .74 to .86. Furthermore, the Gates and Fry lists correlated between .75 and .92 with each other. Evidently, a child who can read phonetically regular words can also read high frequency words, and a child who can identify words in a group situation can do likewise on an individual test. In addition, the Gilmore Accuracy score correlated from .81 to .90 with the Gates Word Pronunciation Test for the various treatments. This would indicate that pronouncing a word in context is closely related to pronouncing a word in isolation. As a further indication of the interrelatedness of reading skills at the first-grade level, it is interesting to note that correlations between rate of reading and the Gates Word Pronunciation Test ranged from .49 to .78, certainly a substantial correlation.

Relationships between teacher, pupil, class characteristics, and achievement

The relationships between various pupil, class, and teacher characteristics and achievement on the Stanford Battery are reported in Table 16. From this table, it is obvious that none of the characteristics is highly related to achievement on any of the Stanford measures. For the class sizes reported in this study, there was no relationship with first-grade achievement. However, there were no very large or very small classes involved in the study. Teacher absence (within the limits of this particular sample) was likewise unrelated to achievement. Teacher experience was positively related to reading achievement with correlations in the neighborhood of .30. However,

Table 10 Intercorrelations of individual outcome measures and selected subtests from the Stanford Achievement Test for the Basal treatment[a]

Variable	1	2	3	4	5	6
1. Stanford word reading		.84	.78	.58	.72	.78
2. Stanford paragraph meaning			.78	.62	.71	.78
3. Gilmore accuracy				.61	.75	.82
4. Gilmore rate					.45	.56
5. Fry word list						.86
6. Gates word list						
Means	20.3	20.0	22.7	56.4	6.9	13.5
Standard deviations	6.87	8.89	11.13	24.31	7.41	7.27

[a] Correlations were calculated within projects and sex and pooled over these units. The N upon which the table is based is 609 coming from 15 projects.

Table 11 Intercorrelations of individual outcome measures and selected subtests from the Stanford Achievement Test for the i.t.a. treatment[a]

Variable	1	2	3	4	5	6
1. Stanford word reading		.84	.80	.59	.69	.83
2. Stanford paragraph meaning			.79	.63	.67	.82
3. Gilmore accuracy				.65	.59	.84
4. Gilmore rate					.44	.67
5. Fry word list						.75
6. Gates word list						
Means	23.8	21.9	24.9	60.5	16.3	18.7
Standard deviations	7.77	10.94	13.57	30.05	10.15	8.53

[a] Correlations were calculated within projects and sex and pooled over these units. The N upon which the table is based is 163 coming from five projects.

Table 12 Intercorrelations of individual outcome measures and selected subtests from the Stanford Achievement Test for the Basal plus Phonics treatment[a]

Variable	1	2	3	4	5	6
1. Stanford word reading		.82	.76	.58	.73	.74
2. Stanford paragraph meaning			.79	.57	.72	.77
3. Gilmore accuracy				.67	.76	.81
4. Gilmore rate					.46	.53
5. Fry word list						.84
6. Gates word list						
Means	23.2	23.2	24.6	58.5	13.3	17.3
Standard deviations	6.52	8.57	10.97	20.62	6.33	7.16

[a] Correlations were calculated within projects and sex and pooled over these units. The N upon which the table is based is 184 coming from three projects.

Table 13 Intercorrelations of individual outcome measures and selected subtests from the Stanford Achievement Test for the Language Experience treatment[a]

Variable	1	2	3	4	5	6
1. Stanford word reading		.84	.78	.49	.81	.86
2. Stanford paragraph meaning			.79	.49	.78	.83
3. Gilmore accuracy				.47	.82	.86
4. Gilmore rate					.45	.49
5. Fry word list						.89
6. Gates word list						
Means	21.4	19.8	21.5	52.7	8.9	13.6
Standard deviations	7.15	9.91	10.89	20.99	8.09	7.60

[a] Correlations were calculated within projects and sex and pooled over these units. The N upon which the table is based is 134 coming from three projects.

Table 14 Intercorrelations of individual outcome measures and selected subtests from the Stanford Achievement Test for the Linguistic treatment[a]

Variable	1	2	3	4	5	6
1. Stanford word reading		.81	.83	.70	.83	.84
2. Stanford paragraph meaning			.85	.70	.79	.84
3. Gilmore accuracy				.81	.88	.90
4. Gilmore rate					.70	.78
5. Fry word list						.92
6. Gates word list						
Means	18.6	15.8	18.4	43.8	8.0	10.7
Standard deviations	7.96	9.33	12.14	27.70	7.19	8.30

[a] Correlations were calculated within projects and sex and pooled over these units. The N upon which the table is based is 146 coming from three projects.

Table 15 Intercorrelations of individual outcome measures and selected subtests from the Stanford Achievement Test for the Phonic/Linguistic treatment[a]

Variable	1	2	3	4	5	6
1. Stanford word reading		.83	.75	.59	.83	.79
2. Stanford paragraph meaning			.78	.64	.75	.72
3. Gilmore accuracy				.65	.83	.84
4. Gilmore rate					.67	.68
5. Fry word list						.87
6. Gates word list						
Means	26.8	24.7	30.1	59.7	18.2	20.5
Standard deviations	6.57	9.64	14.26	25.34	8.76	9.79

[a] Correlations were calculated within projects and sex and pooled over these units. The N upon which the table is based is 94 coming from three projects.

the correlations reported between teacher experience and reading achievement were substantially lower than similar correlations between reading readiness and reading achievement. In general, the younger children did somewhat better in reading than did his older counterpart. Also, in general, the child who attended school regularly did somewhat better than the child who was absent occasionally. However, this relationship was negligible. In summary, it can be said that the teacher characteristics measured in this study were negligibly related to reading success. Furthermore, child age, pupil absence, and class size were related only to a very slight degree.

Data that could not be quantified were also obtained. A great deal of information about school and community characteristics was collected, but much of this was categorical in nature. Information concerning these characteristics within each project is tabled in the appendix to the final report of the Coordinating Center (Bond & Dykstra, 1967).

Evaluation of effectiveness of primary reading programs in the Cooperative Research Program

That part of the analysis which was concerned with evaluating the relative effectiveness of the primary reading programs in the Cooperative Research Program in First-Grade Reading Instruction is presented below. Because not all the various approaches were used in all projects, comparisons could not be made between and among all of them. However, projects which had in common a Basal treatment with another treatment (such as i.t.a.) were grouped together. In this manner, the basal reader treatment was used as a benchmark against which to compare achievement in each of the less traditional nonbasal programs.

General procedures

Data from 15 projects were used in this part of the analysis. These particular 15 projects were included because they used a sample which was considered to be representative of the total population and an experimental program which also was used in another investigation. The establishment of these two criteria eliminated atypical populations such as those comprised of Spanish-speaking youngsters and projects which included a treatment such as individualized reading which was not replicated in any other project.

Six types of instructional materials or methods were used as experimental treatments in more than one project. These six groupings were labeled Initial Teaching Alphabet, Basal plus Phonics, Language Experience, Linguistic, Phonic/Linguistic, and Basal. A listing of the specific materials which comprised each of these major groupings is provided later. Five separate analyses were then performed, each analysis using the basal reader as a control against which to compare progress in other instructional programs. All of the projects which used as experimental treatments both a basal reader approach and an i.t.a. approach, for example, were combined into a single analysis. Similarly, projects were grouped together for analysis if they had in common programs labeled Basal and Language Experience, Basal and Basal plus Phonics, Basal and Linguistic, and Basal and Phonic/Linguistic. It should be emphasized that, for this section of the analysis, methods and materials were placed in categories arbitrarily on the basis of their common characteristics. The purpose was to get some idea of whether or not there was a general superiority of some treatment over several different projects. Major characteristics of each treatment are discussed.

One of the program groupings was labeled the Basal approach. The basal reading program, then, was considered an entity even though the programs of many different publishers were used. The various sets of materials included in this category possess most, if not all, of the following characteristics: (1) Vocabulary is introduced slowly and repeated often. Vocabulary control is based on frequency of usage rather than on regularity of sound-symbol relationships. (2) Phonic analysis is introduced gradually and usually only after some "sight" words have been taught. However, from the beginning the child is encouraged to use such other word recognition skills as context, structural analysis, and picture clues. (3) Emphasis from the beginning is placed not only on word recognition but on comprehension and interpretation of what is read. (4) Silent reading is emphasized early in the program. (5) The various reading skills are introduced and developed systematically. (6) A well-known Basic Reading Series is used as the major instructional tool.

Another method category utilized in this phase of the analysis was labeled i.t.a., or the Initial Teaching Alphabet. This instructional medium purports to simplify the task of learning to read by introducing a novel 44 character alphabet with which to encode the approximately 40 sounds in the English language. In general,

Table 16 Correlations between pupil, class, and teacher characteristics and Stanford Battery[a]

	Word Reading	Paragraph Meaning	Vocabulary	Spelling	Word Study Skills
Class size	−.01	.01	−.03	.02	−.04
Teacher absence	−.07	−.08	−.04	−.01	−.03
Teacher total experience	.27	.32	.24	.34	.32
Teacher first-grade experience	.23	.30	.20	.30	.27
Teacher rating	.14	.13	.10	.22	.19
Child age	−.17	−.17	−.15	−.22	−.20
Pupil absence	−.09	−.05	−.02	−.08	−.09

[a] Correlation coefficients were computed using class means as experimental variables. N = 159.

one symbol is used to represent one sound, thereby making possible more consistent phonic analysis of words. Furthermore, the nature of the alphabet is such that the transition from the use of the i.t.a. to the use of traditional orthography is purported to be a relatively simple task. Two different programs comprised the i.t.a. approach discussed here, but these two programs had in common the unique characteristic of a teaching medium which was quite different from that used in any of the other methods and materials.

A third treatment category was labeled Basal plus Phonics. Each of the treatments in this group was composed of a basal reading series with supplementary phonics materials. The instructional programs, therefore, although somewhat different from project to project, followed a basic philosophy of the basal reader with the addition of a greater phonics emphasis.

A fourth treatment group was labeled Language Experience. A basic element of this instructional method is that the child's own writing serves as a medium of instruction. The child's first stories are dictated to the teacher who acts as the recorder. As soon as he is able, the pupil writes his own stories and shares them with the teacher. During the individual conferences between pupil and teacher, the pupil is helped to recognize the commonality between the words written and spoken and the pupil develops the skills necessary for reading. This approach, then, ordinarily utilizes far fewer highly structured instructional materials than do most instructional programs. In addition, vocabulary control is viewed as being in the language itself and in the language background of each child. The pupil learns to read the words necessary for use in writing. One of the major instructional tasks in this method is to engender a stimulating language environment.

A fifth treatment category was labeled Linguistic. The various materials included in this treatment possess most, if not all, of the following characteristics: (1) There is an early introduction to letters and knowledge of letter names and the ability to recognize letters are considered prerequisite skills for reading instruction. (2) Sound-symbol relationships are taught through careful sequencing of word patterns. Words with high sound-symbol regularity are taught first and the child is led to discover the sound-symbol relationships which exist. In many cases, the child is encouraged to use sound-symbol relationships as the basic word recognition technique by withholding such clues as pictures and word length. (3) In many cases, there is less emphasis on understanding and comprehension in the early stages. Reading is considered a process of translating graphic symbols into sounds and primary attention is paid to helping the child learn the decoding system.

The only "pure" treatment was the Phonic/Linguistic program published by the Lippincott Company. As the name implies, this instructional program has certain characteristics in common with phonic and linguistic programs, as well as with basal programs, but it does not fit well with any of the other treatments. Therefore, the decision was made to recognize this program as a separate method under the category Phonic/Linguistic. It is, in some respects, a linguistically oriented basal program with more demanding pupil expectations.

Description of analysis

The effectiveness of the various reading programs was evaluated in terms of the pupil's end-of-year performance on the five subtests of the Stanford Achievement Test, Primary I Battery, which was administered to all participating pupils after 140 days of instruction in the first grade. In addition, a sample selected from the experimental population was administered the Gilmore Oral Reading Test, the Fry Phonetically Regular Word List, and the Gates Word Pronunciation Test. These tests were individually administered in the testing period immediately following the 140-day instructional period.

Analysis of Stanford Achievement Test scores. The analysis followed a general pattern. For each of the five comparisons (i.t.a. versus Basal, Language Experience versus Basal, Basal plus Phonics versus Basal, Linguistic versus Basal, and Phonic/Linguistic versus Basal), separate means were calculated for males and females within each class on all quantitative variables. The analysis was then conducted using these class means for males and females as the experimental unit, blocking on project, treatment, and sex. This section of the analysis was conducted as if a complete factorial arrangement of treatments had been made. Projects were treated as blocks and the assumption was made that, within each project, treatments were assigned at random to a set of classes. It was assumed that identical treatments were used in each project (within a specified comparison such as i.t.a. versus Basal), thus making it reasonable to test for a general treatment effect over all projects. This portion of the analysis, therefore, gave "across-projects" information.

For each of the five treatment comparisons, an analysis of variance was carried out on the seven premeasures—Murphy-Durrell Phonemes, Murphy-Durrell Letter Names, Murphy-Durrell Learning Rate, Thurstone-Jeffrey Identical Forms, Metropolitan Word Meaning, Metropolitan Listening, and Pintner-Cunningham Primary Test. The analysis of variance on premeasures was designed to indicate those premeasures on which significant differences in performance were found between basal and nonbasal treatments. In this analysis, the Thurstone Pattern Copying Test, which had been administered to all pupils, was not utilized because of its relatively low correlation with the criterion measures and because of the difficulty encountered in scoring the instrument. School, community, and teacher characteristics were not considered in this analysis for two reasons. In the first place, these characteristics, as measured in this investigation, were found to be relatively unrelated to reading achievement. Secondly, many of these characteristics were not quantitative and, in many cases, no ordered relationship existed among categories. As a result, most community characteristics and such teacher characteristics as amount of education could not easily be incorporated as controls in a covariance analysis. All teacher, school, and community characteristics by treatment within project are presented elsewhere (Bond & Dykstra, 1967, Appendix).

The next step was to perform an analysis of covariance using a minimum number of covariates. These were chosen on the basis of their potential for adjusting differences in preinstructional reading-related characteristics. Therefore, the particular premeasures utilized as covariates for an i.t.a. versus Basal comparison might be different from those used for a Language Experience versus Basal comparison.

In each of the five treatment comparisons, a second covariance analysis was also conducted. This covariance analysis utilized all seven premeasures as covariates in order to make pupils in basal and nonbasal treatments as similar as possible in their readiness for reading. This second covariance analysis also had the advantage of being entirely consistent from one treatment comparison to another in that the very same premeasures were used as covariates.

The across-projects covariance analysis of outcome measures was then examined to determine whether or not project by treatment interactions were present. It should be remembered that in this analysis projects were treated as blocks, and analysis of the data ignoring project lines would be meaningful only if no significant project by treatment interactions were found. If such interactions were present, thereby indicating that treatment effects did not operate in the same fashion over all projects, a within-projects analysis was conducted. This within-projects analysis tested for treatment differences within each project but simultaneously for all projects. As a result, all data from all projects involved in a comparison were used to obtain the error term, thus increasing the precision of the experiment. This analysis also followed the pattern of first performing an analysis of variance and then two analyses of covariance, utilizing in turn a selected set of premeasures and the total set of seven premeasures.

The discussion of the method comparisons proceeds in the manner described above. First, the across-projects analysis for each basal versus nonbasal treatment comparison is discussed. Next, the within-projects analysis showing the relative effectiveness of the basal and nonbasal treatments within the projects making up that particular treatment comparison is presented.

Analysis of sample measures. An analysis, similar to the one described for the Stanford Achievement Test results, was conducted on the accuracy and rate scores from the Gilmore Oral Reading Test, as well as on the Fry Phonetically Regular Word Test and the Gates Word Pronunciation Test. Each of these tests was individually administered to a random sample from each treatment within each project. Although these numbers varied from project to project, approximately 20 to 50 pupils were chosen to represent each treatment in each project.

The analysis followed the same steps as those described for Stanford scores. The only difference was that individuals were used as the experimental unit rather than class means based on each sex. With the small numbers involved, it was felt that the use of class means would not have been reasonable. Furthermore, because of consistent project by treatment interactions only the within-projects analysis is reported. The discussion of the within-projects analysis of individual outcome measures follows the discussion of the Stanford data for each of the treatment comparisons.

An illustration of the analysis (Basal versus i.t.a.)

The analysis of the i.t.a. versus Basal treatment comparison demonstrated the technique used for all such comparisons. The discussion of this analysis is presented in greater detail and serves as a model of the analysis used in all cases. The projects in this particular comparison, as well as the numbers of pupils for each treatment and the exact nature of the materials employed, is recorded in Table 17. Four of the five i.t.a. treatments used the Tanyzer-Mazurkiewicz Early-to-Read materials (1964), while one project employed the Downing Readers (Downing, 1963). Although these two sets of materials differ to a considerable extent, the decision was made to pool the data because of the unique similarity regarding the alphabet used for beginning reading instruction. Table 17 also reveals that a variety of basal readers were used in the projects. In fact, in one project, the teachers using basal programs were encouraged to choose any basal series they wished. However, for the purposes of this analysis, the basal treatment was considered to be similar from project to project.

As a first step, an analysis of variance was carried out on the seven premeasures and the five Stanford Reading Achievement Test scores. As was true in all of the basal versus nonbasal treatment comparisons, the experimental unit was a class mean calculated separately

Table 17 Materials and numbers of classes and pupils for Basal vs. i.t.a.

	Fry		Hahn		Hayes		Mazurkiewicz		Tanyzer	
	Classes	Pupils	Classes	Pupils	Classes	Pupils	Classes	Pupils	Classes	Pupils
Numbers										
Basal	6	110	12	276	5	87	17	317	9	228
i.t.a.	7	134	12	255	5	96	15	330	9	240
Materials										
Basal	Allyn-Bacon		Variety		Scott-Foresman		Row-Peterson American Book		Scott-Foresman	
i.t.a.	Early-to-Read		Downing Readers		Early-to-Read		Early-to-Read		Early-to-Read	

for each of the sexes. Projects, treatments, and sex constituted the blocks in the across-projects design. Information pertaining to the analysis of the premeasures is reported in Table 18. Highly reliable project differences were found indicating that pupils in the various projects differed considerably in their readiness for reading. Significant differences were also found favoring girls on five of the seven premeasures. Only one treatment difference was found, that favoring the Basal treatment. Treatment by project interactions was found to be significant on three of the seven premeasures.

The across-projects analysis of variance on the Stanford measures is summarized in Columns A of Table 19. Sex differences favoring girls were found to be significant on four of the five outcome measures. Negligible sex by treatment interactions indicate that boys and girls were not uniquely influenced by either the Basal or i.t.a. treatments. Treatment differences were found to be significant favoring the i.t.a. on the Word Reading Test and Basal pupils on the Spelling Test. The interpretation of differences, however, is clouded by the treatment by project interaction reported to be significant for each of these two measures. Moreover, significant treatment by project interactions were found on the Paragraph Meaning and Word Study Skills variables.

It was hoped that an analysis of covariance might eliminate the project by treatment interactions. The analysis of variance of the premeasures reported in Table 18 was studied to find covariates with the greatest potential for eliminating the interaction. Letter Names was used as a covariate because of the significant treatment by project interaction and because of the large main effects for treatment. Since Phonemes subtest had somewhat the same relationship, it was also included. Columns B of Table 19 report the results of this covariance analysis. The utilization of the Phonemes and Letter Names subtests as covariates reduced the treatment by project interaction on each of the outcome variables, but the same four were still significant. Covariance had the desired effect, but it was not enough to erase the treatment by project interactions. Therefore, the utility of the analysis of treatment differences across projects was still questionable.

The nature of the treatment by project interactions is illustrated in Table 20. It is apparent from this table of unadjusted means that: (1) on the Word Reading test, the only large differences between treatments favored i.t.a.; (2) for the Paragraph Meaning variable, the only large differences favored i.t.a., while small differences in the other projects went both ways; (3) for the Spelling test, the differences were, in general, large but not consistent since project 3 (Hayes) favored i.t.a. while the other differences generally favored Basal; and (4) for the Word Study Skills variable, all differences except those in project 4 (Mazurkiewicz) favored i.t.a., but the differences were of varying amounts.

The analysis of covariance summarized in Columns B of Table 19 also reveals that only one sex difference was recorded, that favoring boys. Adjusting for premeasure differences on the Phonemes and Letter Names tests erased significant differences in reading ability which had been shown to favor girls in the analysis of variance. Project differences, however, were found on each of the five outcome measures.

One last attempt was made to eliminate treatment by project interaction. A covariance analysis using all seven premeasures as covariates was conducted. The result of this covariance analysis is reported in Columns C of Table 19. Very substantial project differences still existed even though pupils' readiness had been controlled. Furthermore, the treatment by project interactions on four of the five variables was still significant. Therefore, the treatment differences found on the Word Reading, Paragraph Meaning, Spelling, and Word Study Skills tests could not be interpreted unambiguously.

As a result of the persistence of the project by treatment interactions, the data were analyzed within

Table 18 Across-projects analysis of variance on premeasures for Basal vs. i.t.a. comparison

Effect	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Sex	13.39***	25.87***	7.88***	4.05**	.34	3.63	12.90***
Treatment	3.38	17.02*	2.22	.77	.41	.08	.11
Sex × Treatment	.63	.34	.00	.24	.65	1.92	.82
Project	17.71††	10.68††	40.52††	10.15††	7.35††	5.84††	11.70††
Sex × Project	.34	.65	.74	.10	.10	.42	.88
Trt × Project	2.94†	5.12††	1.17	.71	3.60††	1.93	1.51
Sex × Trt × Project	.14	.23	.22	.05	.17	.37	.30

* Significant difference favoring Basal at .01 level

** Significant difference favoring females at .05 level

*** Significant difference favoring females at .01 level

† Project difference or interaction significant at .05 level

†† Project difference or interaction significant at .01 level

Note. Numerator *df* = 1 except for all effects involving project in which case numerator *df* = 4. Denominator *df* = 174

Table 19 Across-projects analysis of variance and covariance on Stanford measures for Basal vs. i.t.a. comparison

Effect	Word Reading			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]
Sex	12.53†	.17	.10	19.27†	2.66	2.07	.97	10.72††	8.86††	14.41†	1.86	1.07	9.97†	.06	.15
Treatment	17.08**	29.78**	13.32**	1.19	3.90*	4.65*	.35	.05	.14	13.16***	9.26***	9.44***	3.54	5.16*	6.08*
S × T	.21	.01	.03	.25	.00	.00	.75	.22	.15	1.16	.63	.73	.00	.40	.45
Project	2.51†††	10.36†*	10.83†*	1.62	8.69†*	11.71†*	.80	18.19†*	24.23†*	4.45†*	8.70†*	12.06†*	2.88†††	7.78†*	11.71†*
S × P	.28	.28	.58	.04	.35	.62	.11	.13	.41	.32	.35	.31	.32	.09	.30
T × P	7.11†*	5.00†*	5.21†*	4.20†*	3.34†††	2.86†††	2.21	.83	.21	13.57†*	9.91†*	12.95†*	6.91**	4.07†*	5.52†*
S × T × P	.20	.34	.26	.13	.28	.16	.92	1.51	1.20	.02	.04	.03	.45	.58	.65

[a] Summarizes analysis of variance. Denominator *df* = 174

[b] Summarizes analysis of covariance using Phonemes and Letter Names as covariates. Denominator *df* = 172

[c] Summarizes analysis of covariance using all seven premeasures as covariates. Denominator *df* = 167

* Significant difference favoring i.t.a. at .05 level

** Significant difference favoring i.t.a. at .01 level

*** Significant difference favoring Basal at .01 level

† Significant difference favoring females at .01 level

†† Significant difference favoring males at .01 level

††† Project difference or interaction significant at .05 level

†* Project difference or interaction significant at .01 level

Note. Numerator *df* = 1 for all effects except those involving project which are based on 4 *df*

Table 20 Within-projects unadjusted means on Stanford measures for the Basal vs. i.t.a. comparison

Project	Word Reading		Paragraph Meaning		Vocabulary		Spelling		Word Study Skills	
	Basal	i.t.a.	Basal	i.t.a.	Basal	i.t.a.	Basal	i.t.a.	Basal	i.t.a.
Fry	20.2	20.7	19.9	17.4	23.2	22.3	10.8	7.4	33.9	35.2
Hahn	22.5	24.0	21.8	21.6	21.7	22.0	13.2	11.2	38.5	39.6
Hayes	18.5	24.7	19.3	22.0	22.0	22.2	9.0	14.7	34.2	40.2
Mazurkiewicz	22.0	21.8	21.1	20.1	22.4	20.4	13.6	8.9	39.1	36.1
Tanyzer	17.5	25.3	16.1	23.2	20.5	22.5	10.2	11.9	33.9	42.0

BEST COPY AVAILABLE

Table 21 Within-projects analysis of variance on premeasures for the Basal vs. i.t.a. comparison

Project	Effect	Murphy-Durrell Phonemes	Murphy-Durrell Total Letters	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Fry	Sex × Treatment	.18	.87	.14	.16	.95	.42	1.22
	Treatment	.05	3.02	.78	.17	.25	2.59	1.76
	Sex	3.15	6.57†	5.68†	.46	.01	.81	4.20†
Hahn	Sex × Treatment	.54	.37	.05	.27	.05	.12	.06
	Treatment	1.30	.05	.52	2.92	.35	1.09	.03
	Sex	1.13	2.12	.91	.81	.54	1.02	.54
Hayes	Sex × Treatment	.11	.00	.60	.02	.00	.58	.37
	Treatment	2.61	1.53	.39	.15	1.63	2.69	.16
	Sex	.40	2.37	.03	.44	.07	.27	1.23
Mazurkiewicz	Sex × Treatment	.37	.03	.10	.00	.34	1.19	.28
	Treatment	.46	30.92***	3.96**	.00	10.31***	.41	.88
	Sex	6.11†	7.55††	1.45	.75	.11	1.20	2.53
Tanyzer	Sex × Treatment	.00	.02	.00	.00	.00	1.12	.15
	Treatment	11.70*	1.59	.08	.33	2.31	1.04	2.39
	Sex	3.95†	9.82††	2.76	2.00	.00	2.00	7.86††

* Significant difference favoring i.t.a. at .01 level.
** Significant difference favoring Basal at .05 level.
*** Significant difference favoring Basal at .01 level.
† Significant difference favoring females at .05 level.
†† Significant difference favoring females at .01 level.
Note: All *F* ratios based on 1 and 174 *df*

Table 22 Premeasure means for the Basal vs. i.t.a. comparison

Project	Trt.	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Fry	Basal	23.1	33.4	8.1	15.9	9.2	9.7	40.3
	i.t.a.	22.5	29.3	7.4	15.0	8.8	8.9	38.2
Hahn	Basal	26.9	35.5	8.6	18.7	9.9	9.5	40.0
	i.t.a.	29.2	35.1	9.1	21.6	10.2	9.9	40.2
Hayes	Basal	14.7	24.0	7.0	13.1	7.2	7.8	33.3
	i.t.a.	19.9	27.4	7.6	12.0	8.2	8.8	32.6
Mazurkiewicz	Basal	31.9	37.2	12.5	15.2	9.5	9.0	40.3
	i.t.a.	30.6	28.7	11.5	15.2	8.2	8.8	39.3
Tanyzer	Basal	24.0	36.7	10.9	12.5	8.8	9.7	37.6
	i.t.a.	32.2	34.1	10.7	13.6	9.7	9.3	40.2

each project. This analysis permitted the assessment of the effects of treatment and sex separately for each project. It proceeded in exactly the same fashion as did the across-projects analysis. First, an analysis of variance on the premeasures within each project was carried out. As reported in Table 21, three of the five projects found no treatment differences on any of the seven premeasures. However, within the two remaining projects, significant differences were found in pupil performance on the Phonemes, Letter Names, and Word Meaning subtests. In these projects, the randomization procedure had not succeeded in equalizing prereading ability (as measured by the three subtests) between the basal and nonbasal group. This within-project analysis of premeasures again points out the superiority of girls with respect to prereading ability. The extent of the differences in mean performance on the premeasures between basal and nonbasal groups is illustrated in Table 22 which presents treatment means on each measure for each of the five Basal versus i.t.a. projects.

The next step in the within-projects analysis involved conducting an analysis of variance on the Stanford Achievement measures. This analysis is summarized in Columns A of Table 23. On the Word Reading test, significant differences favoring the i.t.a. treatment were found in two of the five projects. One project recorded a significant difference favoring the i.t.a. treatment on the Paragraph Meaning subtest. On the Vocabulary test, however, the only significant difference favored the Basal treatment. Four of the five projects reported significant differences between treatments on the Spelling subtest, three of these differences favoring the Basal group. This lack of unanimity was further pointed out by the analysis of scores on the Word Study Skills subtest where two projects found significant treatment differences favoring i.t.a., but one project found a significant difference favoring the Basal treatment.

Again, a covariance analysis was run using the Phonemes and Letter Names subtests as covariates. This covariance analysis was conducted within projects although simultaneously for all projects and is summarized in Columns B of Table 23. The results were very similar to those reported for the analysis of variance. One additional treatment difference favoring i.t.a. was found on the Word Reading subtest and the significant difference which had been found favoring the Basal treatment on the Vocabulary subtest was erased, but no changes were reported for the Spelling and Paragraph Meaning subtests. On the Word Study Skills subtest, the covariance analysis erased two of the three significant treatment differences which had been found in the analysis of variance. The use of covariance also tended to eliminate sex differences which had been found to favor girls. Evidently, the superiority of girls in reading achievement at the end of the year could be accounted for by their superiority in prereading capability at the beginning of the year.

The second covariance analysis, utilizing all seven premeasures as covariates, is reported in Columns C of

Table 23 Within-projects analysis of variance and covariance on Stanford measures for the Basal vs. i.t.a. comparison

		Word Reading			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
P[a]	Effect	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]
1	Sex	1.34	.38	.77	2.28	.00	.14	.43	2.14	3.02	2.76	.21	.00	1.64	.05	.19
	Treatment	.10	1.76	2.55	1.29	.38	.15	.48	.00	.16	7.52†	6.56***	5.84***	.38	2.26	3.09
	S × T	.49	.10	.00	.26	.00	.02	4.15*†	4.84*†	4.16*†	.30	.03	.01	.50	.14	.01
2	Sex	.96	.00	.07	4.30††	2.47	3.05	.04	1.31	.52	3.61	1.89	2.03	.69	.00	.11
	Treatment	1.68	1.39	1.92	.01	.36	.25	.07	.03	.01	4.66***	8.37†	8.37†	.43	.05	.17
	S × T	.16	.03	.01	.10	.05	.01	.03	.27	.01	.40	.07	.31	.35	.02	.19
3	Sex	.96	.09	.00	2.52	1.22	.85	.01	1.14	1.67	.85	.12	.02	.33	.02	.17
	Treatment	11.49**	10.29**	17.07**	1.29	.02	1.34	.02	2.01	.20	15.66**	13.56**	21.37**	6.04*	3.46	9.02**
	S × T	.05	.30	.06	.19	.57	.14	.26	.90	.27	.11	.28	.02	.01	.00	.08
4	Sex	6.77†††	.77	1.34	5.67††	.48	.91	.20	5.10†*	2.87	2.43	.00	.03	4.59††	.17	.79
	Treatment	.05	4.89*	9.02*	.55	2.46	1.65	5.57***	.00	.03	35.33†	18.50†	23.17†	5.02***	.45	.78
	S × T	.17	.83	.90	.06	.32	.31	.01	.19	.34	.19	.12	.19	.91	2.58	2.78
5	Sex	3.64	.00	.11	4.64††	.21	.00	.70	2.63	4.07†*	5.97††	1.20	.42	3.97††	.15	.01
	Treatment	32.69**	33.21**	29.92**	15.03**	12.66**	11.41**	3.15	1.17	.43	2.64	.98	.68	19.65**	13.98**	10.92**
	S × T	.11	.14	.13	.16	.19	.18	.01	.08	.14	.31	.35	.35	.02	.01	.03

[a] Projects in numerical order are Fry, Hahn, Mazurkiewicz, and Tanyzer

[b] Summarizes analysis of variance. *F* ratios are based on 1 and 174 *df*

[c] Summarizes covariance using Phonemes and Letter Names as covariates. *F* ratios are based on 1 and 172 *df*

[d] Summarizes covariance using all seven premeasures as covariates. *F* ratios are based on 1 and 167 *df*

* Significant difference favoring i.t.a. at .05 level

** Significant difference favoring i.t.a. at .01 level

*** Significant difference favoring Basal at .05 level

† Significant difference favoring Basal at .01 level

†† Significant difference favoring females at .05 level

††† Significant difference favoring females at .01 level

†* Significant difference favoring males at .05 level

*† Interaction significant at .05 level

BEST COPY AVAILABLE

Table 24 Unadjusted and adjusted Stanford means for the Basal vs. i.t.a. comparison

		Word Reading			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
Project	Treatment	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c
Fry	Basal	20.2	21.2	20.6	19.9	21.0	20.4	23.2	24.0	23.3	10.8	11.4	11.1	33.9	35.3	34.6
	i.t.a.	20.7	22.7	22.4	17.4	20.0	19.8	22.3	24.1	23.7	7.4	8.6	8.7	35.2	37.9	37.3
Hahn	Basal	22.5	22.1	22.0	21.8	21.2	21.2	21.7	21.3	21.0	13.2	12.9	13.0	38.5	38.2	37.9
	i.t.a.	24.0	23.1	23.1	21.6	20.5	20.7	22.0	21.2	21.0	11.2	10.7	10.8	39.6	38.5	38.3
Hayes	Basal	18.5	23.6	23.7	19.3	25.6	25.8	22.0	26.4	26.5	9.0	12.1	12.2	34.2	40.8	40.8
	i.t.a.	24.7	27.7	28.9	22.0	25.9	27.8	22.2	24.9	26.0	14.7	16.5	17.7	40.2	44.2	46.1
Mazurkiewicz	Basal	22.0	20.0	20.0	21.1	18.5	18.2	22.4	20.6	20.7	13.6	12.4	12.1	39.1	36.6	36.5
	i.t.a.	21.8	21.8	21.6	20.1	20.4	19.7	20.4	20.6	20.6	8.9	9.1	8.5	36.1	35.8	35.5
Tanyzer	Basal	17.5	17.7	18.1	16.1	16.0	16.5	20.5	20.4	20.9	10.2	10.1	10.8	33.8	34.2	35.2
	i.t.a.	25.3	23.8	23.8	23.2	21.5	21.5	22.5	21.4	21.7	11.9	11.1	11.2	42.0	40.0	40.1

[a] Reports unadjusted means
[b] Means adjusted for premeasure differences on Phonemes and Letter Names
[c] Means adjusted for all seven covariates

Table 25 Subjects used for the analysis of individual outcome measures for the Basal vs. i.t.a. comparison

Project	Treatment	Males	Females	Total
Fry	Basal	12	15	27
	i.t.a.	23	14	37
Hahn	Basal	24	26	50
	i.t.a.	23	23	46
Hayes	Basal	15	15	30
	i.t.a.	15	15	30
Mazurkiewicz	Basal	12	13	25
	i.t.a.	16	16	32
Tanyzer	Basal	9	8	17
	i.t.a.	8	10	18

Table 23. The utilization of seven premeasures instead of two changed matters very little. Generally speaking, the same conclusions would be drawn from either of these two covariance analyses. In this case, adding covariates beyond the first two served very little purpose.

The unadjusted and adjusted means for the Basal versus i.t.a. groups within each project are reported in Table 24. This table illustrates the actual extent of the difference between the two treatments. Much greater differences in mean performance can be noted for some projects than for others.

Analysis of individual outcome measures. The Gilmore Oral Reading Test, the Fry Phonetically Regular Word Test, and the Gates Word Pronunciation Test were administered individually to a sample from each treatment. The analysis of these test scores followed the same pattern as that described for the Stanford Achievement Test results. However, although both across-projects and within-projects analyses were employed, only the within-projects results are reported. In general, project by treatment interactions were found to exist, thereby making unambiguous interpretation of treatment differences across projects difficult.

Table 25 reports the number of subjects for whom complete information on the premeasures and individual tests was obtained for the Basal versus i.t.a. comparison. The within-projects analysis of variance on the premeasures is summarized in Table 26. Relatively few treatment differences on premeasures are reported. The random selection of pupils from each treatment apparently succeeded quite well in making the two groups of pupils similar in readiness for reading.

The within-projects analysis of variance and covariance of outcome measures is summarized in Table 27. As usual, Columns A summarize the analysis of variance, Columns B summarize an analysis of covariance using a minimum set of covariates, and Columns C report the analysis of covariance using all eight premeasures. (Throughout this section of analysis, eight premeasures are recorded for the individual outcome measures analysis. The Pattern Copying test, which was not used in the analysis of Stanford scores, is used in all of the analyses

Table 26 Within-projects analysis of variance on premeasures for Basal vs. i.t.a. comparison

Project	Effect	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Pattern Copying	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
	Sex	3.03	3.57	1.49	1.30	2.95	.10	.01	6.43†
Fry	Treatment	.05	.72	.83	1.66	.05	.75	.00	1.09
	S × T	1.47	.02	.13	.00	.25	.29	.04	.23
	Sex	.04	.50	1.13	.50	3.74	.00	.13	.33
Hahn	Treatment	.58	.30	.01	2.49	4.05*	.02	3.66	.05
	S × T	.34	1.40	.38	1.12	.37	.08	.47	.00
	Sex	1.18	3.30	.12	.01	1.13	.02	.05	1.12
Hayes	Treatment	5.33*	.25	.37	14.27**	1.20	2.19	.28	2.05
	S × T	.20	.00	.85	1.33	1.41	.24	1.13	1.25
	Sex	2.01	.08	1.73	1.06	.01	2.43	.26	.01
Mazurkiewicz	Treatment	1.04	1.79	.78	.69	.07	1.82	.50	.00
	S × T	.06	.12	.01	.05	.12	.03	.22	.19
	Sex	9.39††	4.46†	1.59	10.58††	.53	.07	3.75	5.66†
Tanyzer	Treatment	4.38*	.76	.68	1.20	.70	2.90	1.82	3.87
	S × T	.30	.03	.35	1.20	.08	.00	1.78	.37

* Significant difference favoring i.t.a. at .05 level
** Significant difference favoring Basal at .01 level
† Significant difference favoring females at .05 level
†† Significant difference favoring females at .01 level
Note: All *F* ratios are based on 1 and 292 *df*

Table 27 Within-projects analysis of variance and covariance on individual outcome measures for Basal vs. i.t.a. comparison

		Gilmore Accuracy			Gilmore Rate			Fry Word List			Gates Word List		
P[a]	Effect	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]
1	Sex	16.07††	12.95††	9.83††	16.39††	13.58††	10.24††	2.27	.46	.13	7.36††	4.15†	2.03
	Treatment	.10	.35	.01	4.70***	5.21***	3.15	5.60*	7.25**	10.09**	.83	.91	2.16
	Sex × Treatment	.00	.36	.20	.38	1.10	.99	.00	.45	.32	.48	2.37	2.18
2	Sex	.44	1.18	.47	.32	.45	.08	.53	.43	1.58	.04	.00	.25
	Treatment	.01	.98	.52	1.11	2.50	1.97	11.75**	10.72**	11.05**	5.52*	4.04*	4.29*
	Sex × Treatment	.00	.00	.27	.30	.14	.00	.05	.00	.11	.00	.01	.35
3	Sex	1.37	.70	.09	9.32††	8.28††	6.65††	2.42	1.55	.80	1.76	1.00	.08
	Treatment	2.84	2.67	3.27	3.21	1.36	1.70	36.62**	36.86**	38.69**	28.38**	32.42**	37.49**
	Sex × Treatment	.82	1.34	1.72	.28	.29	.41	.90	1.17	1.22	.47	.74	.96
4	Sex	.05	.56	.04	1.09	.41	1.34	.02	1.07	.35	.00	1.12	.50
	Treatment	1.14	2.74	.76	1.43	.81	2.57	.83	.37	1.96	.00	.13	.04
	Sex × Treatment	.45	.77	1.33	.14	.19	.37	.13	.25	.34	.22	.24	.15
5	Sex	1.85	.76	.61	.80	.12	.03	5.18†	.24	.53	1.18	1.45	1.56
	Treatment	9.08**	4.78*	8.97**	.44	.02	.09	48.06**	45.50**	56.38**	28.83**	25.27**	35.10**
	Sex × Treatment	1.55	2.26	2.14	1.41	1.49	1.13	1.28	1.93	2.06	.19	.24	.09

[a] Projects in numerical order are Fry, Hahn, Hayes, Mazurkiewicz, and Tanyzer
[b] Summarizes analysis of variance. *F* ratios are based on 1 and 292 *df*
[c] Summarizes covariance using Phonemes, Pattern Copying, and Listening as covariates. *F* ratios are based on 1 and 289 *df*
[d] Summarizes covariance using all eight premeasures as covariates. *F* ratios are based on 1 and 284 *df*
* Significant difference favoring i.t.a. at .05 level
** Significant difference favoring i.t.a. at .01 level
*** Significant difference favoring Basal at .05 level
† Significant difference favoring females at .05 level
†† Significant difference favoring females at .01 level

BEST COPY AVAILABLE

of individual outcome measures.) As reported in Column C, there were no differences between the i.t.a. subjects and the Basal subjects on the accuracy score of the Gilmore Oral Reading Test in four of the five projects. The one project which reported a significant difference between treatments indicated that these differences favored the i.t.a. group. In terms of reading rate, none of the five projects found significant differences between treatments. However, there were pronounced differences on the two Word Recognition tests. Four of the five projects reported significant differences on the Fry Word List, all of which favored the i.t.a. treatment. Three of the five projects found significant differences in favor of the i.t.a. treatment on the Gates test. This finding supports the results of the analysis of treatment differences on the Stanford Word Reading test where differences were also found to favor the i.t.a. approach.

The actual unadjusted and adjusted means for the i.t.a. and Basal treatments are reported in Table 28. The differences in mean performance on the Fry and Gates word lists are often quite striking.

Summary of Basal versus i.t.a. comparisons. The i.t.a. and Basal approaches were of approximately equal effectiveness in terms of pupils' achievement on the Paragraph Meaning test. However, the i.t.a. treatment produced superior word recognition abilities as measured by the Word Reading subtest of the Stanford and the Fry and Gates word recognition lists. Evidence concerning the spelling ability of pupils in the two groups was inconclusive. The Basal subjects were superior in spelling ability in three projects, but the i.t.a. subjects were superior in a fourth project. Furthermore, no differences were found between treatments in reading accuracy and rate as measured by the Gilmore Oral Reading Test.

In interpreting the results of the i.t.a. versus Basal comparisons, it should be pointed out that all testing was done in traditional orthography. Furthermore, children were judged to spell a word correctly only if they spelled it correctly in the traditional sense. No credit was given for spelling a word correctly according to the rules of i.t.a. Varying proportions of children in each of the projects were still receiving instruction in i.t.a. at the time of testing and had not made formal transition to traditional orthography. Therefore, many of the pupils were asked to take a test in an orthography which they had not used during their instruction in reading.

Basal versus Basal plus Phonics comparisons

The Basal plus Phonics versus Basal treatment comparison was analyzed in a manner similar to that outlined for the i.t.a. versus Basal comparison. However, in this section, as well as the sections which follow, the analysis is presented in much less detail. Information about projects which were involved in the Basal versus Basal plus Phonics comparison is provided in Table 29. Four projects with varying numbers of students had in common a Basal treatment and a treatment which could be considered a basal reading program with supplementary phonics. The nature of the materials is also recorded in Table 29. Two of the four projects used exactly the same Basal plus Phonics program. Here again, however, any difference among programs within either the Basal or Basal plus Phonics treatments was ignored.

Table 28 Unadjusted and adjusted means on individual outcome measures for the Basal vs. i.t.a. comparison

		Gilmore Accuracy			Gilmore Rate			Fry Word List			Gates Word List		
Project	Treatment	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]
Fry	Basal	23.9	27.6	25.6	59.8	62.9	58.5	5.4	7.5	6.8	12.3	14.4	13.5
	i.t.a.	22.9	26.1	25.8	44.6	47.9	47.3	10.8	12.7	12.8	14.0	15.9	15.8
Hahn	Basal	24.9	24.5	23.3	67.6	66.4	63.2	10.4	10.0	9.2	15.6	15.3	14.5
	i.t.a.	24.6	22.4	21.9	61.7	57.9	55.9	16.7	15.2	14.4	19.3	18.0	17.1
Hayes	Basal	19.2	25.2	25.9	65.2	75.7	77.4	3.8	8.3	8.9	11.4	15.3	15.7
	i.t.a.	24.7	29.8	30.7	78.0	83.9	86.2	17.8	20.9	21.4	21.9	25.0	25.6
Mazurkiewicz	Basal	25.1	19.6	19.7	56.6	52.3	53.1	13.5	10.5	10.4	17.1	13.9	14.4
	i.t.a.	21.5	15.0	17.4	65.5	58.6	63.9	15.7	11.8	13.3	17.3	13.3	14.7
Tanyzer	Basal	23.4	23.3	22.8	45.9	45.8	46.2	3.8	3.8	3.3	9.9	9.9	9.9
	i.t.a.	36.2	31.1	32.8	52.1	44.6	48.7	24.8	21.4	22.4	23.8	20.6	22.1

[a] Reports unadjusted means

[b] Means adjusted for Phonemes, Pattern Copying, and Listening

[c] Means adjusted for all eight premeasures

BEST COPY AVAILABLE

Table 29 Materials and numbers of classes and pupils for Basal vs. Basal plus Phonics

	Bordeaux		Hayes		Manning		Murphy	
	Classes	Pupils	Classes	Pupils	Classes	Pupils	Classes	Pupils
Numbers								
Basal	5	111	5	87	13	310	10	214
Basal plus Phonics	5	119	5	103	12	200	20	518
Materials								
Basal	Scott-Foresman		Scott-Foresman		Ginn		Scott-Foresman	
Basal plus Phonics	Scott-Foresman + Speech to Print		Scott-Foresman + Phonic Word Power		Ginn + Special Phonic Exercises		Scott-Foresman + Speech to Print	

Table 30 Across-projects analysis of variance on premeasures for Basal vs. Basal plus Phonics comparison

Effect	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Sex	4.90**	5.22**	5.19**	2.26	.63	.12	5.65**
Treatment	1.19	4.94*	.51	.00	1.56	6.39*	.02
Sex × Treatment	.06	.12	.49	.01	.02	.03	.06
Project	12.40††	8.23††	14.09††	12.85††	3.68†	27.29††	19.21††
Sex × Project	.03	.05	.40	.20	.39	.47	.09
Treatment × Project	.73	4.39††	1.62	1.74	2.19	.26	1.05
Sex × Trt. × Proj.	.60	.05	.71	.13	.06	.73	.28

* Significant difference favoring Basal plus Phonics at .05 level

** Significant difference favoring females at .05 level

† Project difference or interaction significant at .05 level

†† Project difference or interaction significant at .01 level

Note Numerator *df* = 1 except for all effects involving project in which case numerator *df* = 3. Denominator *df* = 134

The first step again involved an analysis of variance on the premeasures and Stanford tests blocking on sex, treatment, and project. The across-projects analysis of variance on the premeasures is summarized in Table 30. Highly significant project differences were found on each of the premeasures. A number of sex differences also showed girls predominating. Only two treatment differences were reported, both of these favoring the Basal plus Phonics subjects. Furthermore, only one treatment by project interaction was found to be significant.

The across-projects analysis of variance and covariance on the Stanford measures is reported in Table 31. Significant treatment differences on all five reading achievement measures were found to favor the Basal plus Phonics approach. Furthermore, sex differences were found to be significant and in favor of girls on three of the five outcome measures even when the scores were adjusted for premeasure differences. Similarly, highly reliable project differences were found on all measures in both covariance analyses, again indicating that projects differed on important reading-related characteristics other than pupil readiness. Perhaps the most interesting information in Table 31 is that regarding the treatment by project interactions. In the covariance analyses, no treatment by project interactions were found to be significant. Apparently, the Basal plus Phonics and Basal treatments were operating in the same fashion within each project. This analysis graphically illustrates the superiority of the Basal plus Phonics approach over the Basal alone.

Despite the absence of treatment by project interactions, in the interests of consistency, a within-projects analysis was also employed. The analysis of variance on premeasures is reported in Table 32. Except for one project, no treatment difference on any premeasure was found to exist. The actual premeasure means by experimental treatment are reported in Table 33. The similarity of treatment means on the various premeasures within projects demonstrates the effectiveness of the random assignment of pupils or classes to treatment.

BEST COPY AVAILABLE

Table 31 Across-projects analysis of variance and covariance on Stanford measures for Basal vs. Basal plus Phonics comparison

	Word Recognition			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
Effect	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]
Sex	6.13**	2.74	4.68**	16.80***	13.32***	16.10***	.53	.18	.00	6.93**	3.42	4.40**	5.21**	2.57	3.15
Treatment	20.86*	23.43*	18.30*	15.89*	16.43*	12.98*	14.16*	12.15*	8.66*	22.37*	16.19*	13.53*	24.47*	24.10*	21.94*
S × T	.08	.28	.13	.39	.40	.82	.00	.01	.08	.01	.00	.04	.00	.00	.08
Project	14.82††	29.17††	13.85††	9.89††	22.30††	7.28††	8.77††	17.45††	8.02††	11.04††	17.80††	3.54†	19.38††	22.65††	13.88††
S × P	.18	.14	.09	.43	.51	.51	.21	.29	.34	.10	.12	.17	.16	.09	.11
T × P	1.25	.07	.21	1.26	.86	.76	3.45†	1.95	2.11	2.52	.90	.75	1.77	.60	.52
S × T × P	.30	.37	.41	.35	.37	.52	.73	.81	1.18	.30	.25	.21	.38	.24	.10

[a] Summarizes analysis of variance. Numerator *df* = 1 for all effects except those involving project which are based on 3 *df*. Denominator *df* = 134

[b] Summarizes covariance using Letter Names and Listening as covariates. Numerator *df* = 1 for all except those involving project which are based on 3 *df*. Denominator *df* = 132

[c] Summarizes covariance using all seven premeasures as covariates. Numerator *df* = 1 for all effects except those involving project which are based on 3 *df*. Denominator *df* = 127

* Significant difference favoring Basal plus Phonics at .01 level

** Significant difference favoring females at .05 level

*** Significant difference favoring females at .01 level

† Project difference or interaction significant at .05 level

†† Project difference or interaction significant at .01 level

Table 32 Within-projects analysis of variance on premeasures for the Basal vs. Basal plus Phonics comparison

Project	Effect	Murphy-Durrell Phonemes	Murphy-Durrell Total Letters	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
	Sex × treatment	1.43	.00	1.91	.23	.21	1.98	.39
Bordeaux	Treatment	.01	.04	1.57	2.96	.22	.32	.85
	Sex	.37	.52	2.46	.43	.01	.47	.23
	Sex × Treatment	.39	.00	.42	.04	.01	.01	.03
Hayes	Treatment	.01	1.28	.01	1.24	.80	.04	.95
	Sex	.85	1.17	2.12	1.26	.04	.09	1.50
	Sex × Treatment	.01	.00	.04	.12	.02	.26	.41
Manning	Treatment	4.01*	16.45**	3.93*	.03	5.86*	.12	1.10
	Sex	1.40	1.11	.69	.97	1.70	.52	1.68
	Sex × Treatment	.04	.24	.18	.02	.01	.02	.06
Murphy	Treatment	.04	.26	1.75	1.93	.44	1.45	.29
	Sex	2.35	2.60	1.21	.21	.00	.42	2.35

* Significant difference favoring Basal plus Phonics at .05 level

** Significant difference favoring Basal plus Phonics at .01 level

Note. All *F* ratios based on 1 and 132 *df*

The within-projects analysis of variance and covariance on the Stanford measures is summarized in Table 34. It is evident that the superiority of the Basal plus Phonics treatment was not as clearcut in this within-projects analysis as it had been in the across-projects analysis. In the covariance analysis reported in Columns C, none of the four projects showed significant treatment differences on the Word Reading variable. Only one significant difference was found for the Paragraph Meaning subtest, the Vocabulary subtest, and the Spelling subtest, while two significant differences were found on the Word Study Skills test. All significant differences favored the Basal plus Phonics approach, but the superiority of this program was not nearly so apparent in this type of analysis.

The unadjusted and adjusted Stanford means for the Basal versus Basal plus Phonics comparison are reported in Table 35. The table indicates that practically all of the mean differences favored the Basal plus Phonics treatment. In the within-projects analysis, many of these differences were not statistically significant. However, when

BEST COPY AVAILABLE

the data from the four projects were pooled in the across-projects analysis, the resulting differences did prove to be significant, favoring the Basal plus Phonics approach.

Analysis of individual outcome measures. The number of subjects who comprised the Basal and Basal plus Phonics sample groups for the individual analysis is reported in Table 36. The number of subjects chosen for the individual tests varied considerably from project to project. The within-projects analysis of variance on premeasures for the sample subjects is reported in Table 37. Only one significant sex difference was found and relatively few treatment differences were reported. The analysis of variance and covariance on the individual outcome measures is reported in Table 38. The covariance analysis reported in Columns C found no differences in rate of reading between the two treatments. One of the four projects found a significant difference favoring the Basal group in reading accuracy. Differences on the Fry and Gates word lists tended to favor the Basal plus Phonics group with three such differences reaching statistical significance. The actual unadjusted and adjusted treatment means are reported in Table 39. The Basal

Table 33 Premeasure means for the Basal vs. Basal plus Phonics comparison

Project	Treatment	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Bordeaux	Basal	21.2	31.6	9.6	14.2	8.0	9.9	43.7
	B + P	21.5	30.8	8.3	18.1	8.3	9.5	41.6
Hayes	Basal	14.7	24.0	7.0	13.0	7.2	7.8	33.3
	B + P	15.0	19.9	7.1	10.6	6.6	8.0	31.1
Manning	Basal	24.5	28.0	9.3	15.0	6.8	9.5	32.1
	B + P	28.8	37.4	10.6	14.7	7.8	9.3	33.7
Murphy	Basal	21.2	29.9	6.9	8.7	8.1	7.5	33.7
	B + P	21.6	31.1	7.7	10.6	7.8	7.0	34.4

Table 34 Within-projects analysis of variance and covariance on Stanford measures for the Basal vs. Basal plus Phonics comparison

	Word Recognition			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
P[a]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]
1 Sex	2.10	1.48	1.69	4.70†	4.33†	4.74†	.88	.36	.54	1.58	1.06	1.28	.74	.29	.38
Treatment	.18	.71	.15	.01	.02	.07	.16	.50	.16	.27	.48	1.12	.00	.04	.16
S × T	.21	.02	.01	.03	.04	.19	.34	.16	.00	.47	.41	.15	1.02	.62	.19
2 Sex	.77	.14	.35	4.58†	3.87	4.53†	.21	.01	.00	1.60	.82	1.02	.54	.13	.12
Treatment	.32	1.82	2.14	.01	.41	.49	2.47	1.70	1.93	.81	2.08	2.15	.26	.91	1.08
S × T	.09	.15	.28	.00	.00	.06	.01	.01	.23	.05	.06	.01	.08	.10	.01
3 Sex	.96	.69	1.61	3.73	3.98†	5.61†	.04	.04	.01	2.18	1.45	2.14	.71	.51	.71
Treatment	12.13**	3.76	2.52	7.64**	1.68	1.14	11.29**	3.01	2.64	19.77**	8.38**	7.86**	14.23**	6.66*	7.45**
S × T	.41	1.02	.69	.00	.01	.00	.39	.78	.55	.01	.03	.00	.04	.01	.04
4 Sex	2.81	.95	1.76	4.97†	2.90	3.95†	.03	.64	.34	1.85	.57	.80	3.69	1.96	2.59
Treatment	2.20	3.75	3.72	5.65*	8.58**	8.28**	4.45*	6.07*	6.06*	1.10	1.00	.92	3.76	5.13*	4.56*
S × T	.30	.21	.39	1.51	1.66	2.47	1.45	1.46	2.93	.40	.25	.55	.02	.00	.11

a Projects in numerical order are Bordeaux, Hayes, Manning, and Murphy

b Summarizes analysis of variance. All *F* ratios are based on 1 and 134 *df*.

c Summarizes covariance using Letter Names and Listening as covariates. All *F* ratios are based on 1 and 132 *df*.

d Summarizes covariance using all seven premeasures as covariates. All *F* ratios are based on 1 and 127 *df*.

* Significant difference favoring Basal plus Phonics at .05 level

** Significant difference favoring Basal plus Phonics at .01 level

† Significant difference favoring females at .05 level

BEST COPY AVAILABLE

Table 35 Unadjusted and adjusted Stanford means for the Basal vs. Basal plus Phonics comparison

		Word Reading			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
Project	Treatment	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c
Bordeaux	Basal	19.0	17.6	18.5	20.5	18.9	18.4	19.8	18.9	19.2	9.8	9.3	8.4	32.5	30.9	29.6
	B + P	19.8	18.8	19.0	20.3	19.2	18.9	20.5	19.8	19.7	10.7	10.4	10.0	32.5	31.4	30.5
Hayes	Basal	18.5	20.3	20.2	19.3	21.2	21.0	22.0	23.5	23.6	9.0	10.2	10.2	34.2	35.9	35.9
	B + P	19.6	22.3	22.3	19.5	22.4	22.3	19.4	21.6	21.8	10.6	11.5	12.4	35.4	38.0	38.0
Manning	Basal	16.0	15.7	16.2	14.7	14.3	15.8	17.4	17.3	18.3	4.9	5.1	6.5	27.5	27.1	29.1
	B + P	20.1	17.7	17.8	18.7	16.0	17.1	20.9	19.0	19.7	9.7	8.2	9.6	33.4	31.0	32.9
Murphy	Basal	22.4	23.0	22.5	20.2	20.9	20.1	21.6	22.0	21.3	11.1	11.3	10.4	37.0	37.8	36.7
	B + P	24.0	24.7	24.2	23.5	24.3	23.4	23.7	24.1	23.2	12.2	12.3	11.4	39.9	40.8	39.4

[a] Reports unadjusted means.
[b] Means adjusted for premeasure differences on Letter Names and Listening.
[c] Means adjusted for all seven covariates.

Table 36 Subjects used for the analysis of individual outcome measures for the Basal vs. Basal plus Phonics comparison

Projects	Treatment	Males	Females	Total
Bordeaux	Basal	9	10	19
	B + P	10	10	20
Hayes	Basal	15	15	30
	B + P	15	15	30
Manning	Basal	29	35	64
	B + P	30	26	56
Murphy	Basal	26	22	48
	B + P	57	41	98

plus Phonics treatment tended to surpass the Basal treatment in performance on the two word lists, but no trend was apparent on the Gilmore Oral Reading Test.

Summary of Basal versus Basal plus Phonics comparison. In general, basal program accompanied by supplementary phonics materials led to significantly greater achievement in reading than did basal materials alone. This superiority was especially pronounced in mean performance on the Stanford Achievement Test and the Fry and Gates word recognition tests. Practically all differences on these measures favored the Basal plus Phonics group (particularly in the across-projects analysis), even though some of the differences failed to reach statistical significance. No differences in rate or accuracy of reading were found between the two treatments.

Basal versus Language Experience comparisons

Four projects had as experimental treatments both the Language Experience approach and the Basal reader approach. Information about the numbers of classes involved and the nature of the Basal readers is provided in Table 40. For purposes of this analysis, the Basal approach was considered a single method, even though a variety of basal readers were employed. Likewise, the Language Experience approach differed considerably in its implementation from one project to another. However, the Language Experience approaches had more similarities than differences and, therefore, were considered to constitute a single treatment.

Again, the first step was to carry out an analysis of variance on both the premeasures and the postmeasures blocking on treatment, sex, and project. The across-projects analysis of variance on premeasures is reported in Table 41. Among the interesting results of this analysis were the superiority of girls on the premeasures, the highly reliable project differences on six of the seven premeasures, the treatment differences on four of the seven premeasures, and the absence of treatment by project interactions on six of the seven premeasures.

The analysis of variance and covariance across projects on the Stanford measures is reported in Table 42. The differences in mean achievement among projects is graphically illustrated by the highly reliable *F* ratios reported for the projects' main effect in the two covariance analyses. It is also apparent from Table 42 that sex differences in achievement tend to disappear when differences in premeasure capability are taken into account. Treatment differences in the analysis of variance, as reported in Columns A, tended to be negligible. However,

Table 37 Within-projects analysis of variance on premeasures for Basal vs. Basal plus Phonics comparison

Project	Effect	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Pattern Copying	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
	Sex	.28	.24	.22	.11	.39	4.42††	.05	.95
Manning	Treatment	3.88	7.98**	7.65**	.25	.40	5.16*	2.17	.02
	S × T	.14	.26	.20	1.90	.50	.24	.70	.04
	Sex	.25	.41	.81	.01	3.11	.26	.01	3.18
Hayes	Treatment	.00	1.49	.54	12.31†	3.11	.49	.46	4.55***
	S × T	.01	.84	.16	.74	.31	1.36	1.13	.07
	Sex	.30	.08	.32	.70	.99	1.70	.96	.42
Murphy	Treatment	.49	.65	3.05	11.79**	.39	.16	.08	2.29
	S × T	1.07	.00	1.71	3.56	.04	.80	1.15	.09
	Sex	.08	.84	.63	.33	1.38	.05	.76	.20
Bordeaux	Treatment	.32	2.47	.01	.03	4.93*	4.67*	.01	.23
	S × T	.50	.12	.14	.35	.09	.06	1.93	.08

* Significant difference favoring Basal plus Phonics at .05 level

** Significant difference favoring Basal plus Phonics at .01 level

*** Significant difference favoring Basal at .05 level

† Significant difference favoring Basal at .01 level

†† Significant difference favoring males at .05 level

Note. All *F* ratios are based on 1 and 349 *df*

Table 38 Within-projects analysis of variance and covariance on individual outcome measures for Basal vs. Basal plus Phonics comparison

		Gilmore Accuracy			Gilmore Rate			Fry Word List			Gates Word List		
P[a]	Effect	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]
1	Sex	1.14	.63	.78	2.42	1.99	2.12	.01	.37	.34	.49	.08	.13
	Treatment	1.53	.03	1.40	.14	.25	.00	4.29*	1.87	3.87	.90	.01	.45
	Sex × Treatment	.75	1.22	.85	.58	.75	.57	1.00	1.46	1.00	.73	1.30	.87
2	Sex	.39	.09	.07	2.14	1.69	1.32	.00	.18	.53	.25	.02	.07
	Treatment	.05	2.73	2.81	.43	.00	.01	.08	1.96	1.55	.29	.30	.20
	Sex × Treatment	.10	.58	.62	2.44	2.46	2.89	.75	.38	.45	.04	.04	.05
3	Sex	2.59	4.60†	2.80	1.55	2.43	2.11	.19	.25	.00	.65	.97	.25
	Treatment	4.92*	.43	1.44	4.05*	.69	.68	6.37*	1.55	2.32	9.18**	2.72	4.58*
	Sex × Treatment	.13	.03	.07	1.62	3.06	2.68	.23	.01	.16	.19	.00	.16
4	Sex	1.13	2.58	1.38	1.64	2.60	2.15	.00	.03	.16	.00	.03	.09
	Treatment	.82	3.98***	4.56***	1.33	2.77	3.08	37.25**	38.54**	45.12**	15.48**	15.62**	17.46**
	Sex × Treatment	.21	.23	1.47	.46	.49	.08	.79	.83	1.74	1.86	2.56	4.92††

[a] Projects in numerical order are Bordeaux, Hayes, Manning, and Murphy

[b] Summarizes analysis of variance. All *F* ratios are based on 1 and 349 *df*

[c] Summarizes covariance using Letter Names, Pattern Copying, and Meaning as covariates. All *F* ratios are based on 1 and 346 *df*

[d] Summarizes covariance using all eight premeasures as covariates. All *F* ratios are based on 1 and 341 *df*

* Significant difference favoring Basal plus Phonics at .05 level

** Significant difference favoring Basal plus Phonics at .01 level

*** Significant difference favoring Basal at .05 level

† Significant difference favoring females at .05 level

†† Interaction significant at .05 level

Table 39 Unadjusted and adjusted means on individual outcome measures for the Basal vs. Basal plus Phonics comparison

		Gilmore Accuracy			Gilmore Rate			Fry Word List			Gates Word List		
Project	Treatment	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]
Bordeaux	Basal	21.4	24.2	19.6	58.4	62.6	58.5	5.0	6.4	4.8	10.9	12.6	10.3
	B + P	25.8	24.7	22.8	61.2	59.2	58.7	9.4	8.9	8.3	13.1	12.4	11.5
Hayes	Basal	19.2	22.5	22.7	65.2	70.1	71.1	3.8	5.5	6.1	11.4	13.4	13.6
	B + P	19.8	26.4	26.4	61.3	70.3	70.7	4.3	7.6	7.9	10.4	14.2	14.2
Manning	Basal	14.8	14.9	14.5	52.0	52.8	52.0	2.1	2.1	1.9	8.7	8.8	8.6
	B + P	19.3	16.0	16.4	60.5	56.0	55.2	5.1	3.4	3.4	12.5	10.5	10.7
Murphy	Basal	30.9	30.8	31.6	61.3	60.5	61.8	13.7	13.7	13.6	17.3	17.2	17.6
	B + P	29.1	27.6	28.4	56.6	54.3	55.2	20.7	20.0	20.3	22.1	21.2	21.7

[a] Reports unadjusted means
[b] Means adjusted for Letter Names, Pattern Copying, and Meaning
[c] Means adjusted for all eight premeasures

Table 40 Materials and numbers of classes and pupils for Basal vs. Language Experience

	Cleland		Hahn		Kendrick		Stauffer	
	Classes	Pupils	Classes	Pupils	Classes	Pupils	Classes	Pupils
Numbers								
Basal	12	376	12	276	27	652	10	219
Language Experience	11	287	12	269	27	637	10	238
Materials								
Basal	Scott-Foresman		Variety		Ginn		Variety	

Table 41 Across-projects analysis of variance on premeasures for Basal vs. Language Experience comparison

Effect	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner Cunningham IQ
Sex	11.31†	10.06†	4.23***	8.82†	5.70††	.32	7.16†
Treatment	9.92**	1.54	4.18*	11.99**	5.29*	.61	1.89
Sex × Treatment	.76	.03	.06	.55	.07	.06	.45
Project	36.37†††	18.54†††	18.17†††	1.64	16.43†††	7.23†††	7.95†††
Sex × Project	.06	.41	.40	.11	.10	.33	.23
Trt. × Project	11.17†††	1.20	.37	2.03	2.19	1.42	1.27
Sex × Trt. × Proj.	.11	.20	.08	.27	.16	.56	.45

* Significant difference favoring Language Experience at .05 level
** Significant difference favoring Basal at .01 level
*** Significant difference favoring females at .05 level
† Significant difference favoring females at .01 level
†† Significant difference favoring males at .05 level
††† Project difference or interaction significant at .01 level
Note. Numerator *df* = 1, except for all effects involving project in which numerator *df* = 3. Denominator *df* = 226.

BEST COPY AVAILABLE

Table 42 Across-projects analysis of variance and covariance on Stanford measures for Basal vs. Language Experience comparison

	Word Reading			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
Effect	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]
Sex	7.57**	.00	.81	18.25**	5.56**	6.95**	.12	7.44†	3.10	12.42**	2.11	2.98	8.16**	.06	.98
Treatment	2.53	35.92*	8.40*	1.44	2.28	.34	2.09	24.66*	1.49	.03	6.77*	.06	.03	14.60*	.21
S × T	.11	.28	.16	.00	.00	.03	.51	.98	.42	.06	.08	.02	.09	.21	.04
Project	25.81†††	68.28†††	52.44†††	30.44†††	75.57†††	62.78†††	11.65†††	22.91†††	20.15†††	30.87†††	60.91†††	49.06†††	14.88†††	53.87†††	43.27†††
S × P	.10	.40	.26	.13	.25	.33	.10	.25	.27	.01	.01	.03	.11	.28	.13
T × P	2.23	1.43	2.95††	5.49††	2.57	5.26†††	.56	3.35††	4.19†††	1.40	.74	.88	1.04	2.64	2.86††
S × T × P	.20	.22	.23	.28	.27	.34	.45	.45	.27	.17	.14	.15	.33	.31	.25

[a] Summarizes analysis of variance. Numerator *df* = 1 for all except those involving project which are based on 3 *df*. Denominator *df* = 226.

[b] Summarizes covariance using Phonemes and Identical Forms as covariates. Numerator *df* = 1 for all effects except those involving project which are based on 3 *df*. Denominator *df* = 224.

[c] Summarizes covariance using all seven premeasures as covariates. Numerator *df* = 1 for all effects except those involving project which are based on 3 *df*. Denominator *df* = 219.

* Significant difference favoring Language Experience at .01 level.

** Significant difference favoring females at .01 level.

*** Significant difference favoring males at .05 level.

† Significant difference favoring males at .01 level.

†† Project difference or interaction significant at .05 level.

††† Project difference or interaction significant at .01 level.

Table 43 Within-projects analysis of variance on premeasures for the Basal vs. Language Experience comparison

Project	Effect	Murphy-Durrell Phonemes	Murphy-Durrell Total Letters	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
	Sex × Treatment	.36	.25	.09	.69	.11	.71	.02
Cleland	Treatment	4.49***	.86	.03	3.33	1.18	.46	.15
	Sex	2.34	2.14	.06	.89	.38	.01	1.27
	Sex × Treatment	.00	.01	.06	.02	.27	.07	.06
Hahn	Treatment	.14	.17	.29	.11	.42	.85	.05
	Sex	3.18	3.56	.64	1.50	1.10	.87	.41
	Sex × Treatment	.02	.15	.15	.00	.00	.50	.29
Kendrick	Treatment	30.22†	4.04*	3.99*	2.74	8.69**	2.65	5.32*
	Sex	3.87††	1.85	4.58††	4.08††	2.78	.01	3.22
	Sex × Treatment	.03	.22	.01	.67	.18	.45	1.42
Stauffer	Treatment	8.98**	.09	1.11	11.80†	1.44	.87	.10
	Sex	2.09	3.73	.16	2.67	1.74	.42	2.96

* Significant difference favoring Language Experience at .05 level.

** Significant difference favoring Language Experience at .01 level.

*** Significant difference favoring Basal at .05 level.

† Significant difference favoring Basal at .01 level.

†† Significant difference favoring females at .05 level.

Note: All *F* rations are based on 1 and 226 *df*.

BEST COPY AVAILABLE

the analysis of covariance reported in Columns B, in which Phonemes and Identical Forms are used as covariates, found significant differences favoring the Language Experience approach on the Word Reading test, the Vocabulary test, the Spelling test, and the Word Study Skills test. Strangely enough, these treatment differences were erased for all but the Word Reading test when covariance analysis was performed using all seven premeasures as covariates. This unusual set of events is probably a result of the peculiar nature of the treatment differences on premeasures as reported in the across-projects analysis in Table 41. Significant treatment differences were found in the across-projects analysis for four of the seven premeasures. In two cases, the difference favored the Basal group, but, in the other two cases, the difference favored the Language Experience approach. At any rate, one would likely draw different conclusions about the effectiveness of Language Experience and Basal approaches depending on which of the analyses he considers. Quite different results were obtained by the analysis of variance reported in Columns A and the analysis of covariance reported in Columns B. Similarly, quite different results were found between the covariance analyses reported in Columns B and C.

One further finding of note in the across-projects analysis of the Stanford concerns the treatment by project interactions reported in Table 42. The covariance analysis using all seven premeasures as covariates found significant treatment by project interaction on four of the five Stanford measures. This finding made it necessary to look to the within-projects analysis for an assessment of the relative effectiveness of the Basal and Language Experience programs. However, it would be of interest to find the reason for the project by treatment interactions. The means reported for the Stanford tests and the treatments within each project on Table 46 (Columns C) reveal that the Language Experience subjects in each of the projects were superior on the Word Reading test.

Table 44 Pre-measure means for the Basal vs. Language Experience comparison

Project	Treatment	Murphy-Durrell Phonemes	Murphy-Durrell Total Letters	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Cleland	Basal	24.7	34.9	10.0	19.4	8.7	9.2	37.1
	Language Experience	20.2	32.9	10.1	16.1	8.2	8.9	36.5
Hahn	Basal	26.9	35.5	8.6	18.7	9.9	9.5	40.0
	Language Experience	27.6	36.3	9.0	18.1	10.2	9.2	40.3
Kendrick	Basal	34.2	33.9	9.8	17.8	8.4	9.4	37.5
	Language Experience	26.6	36.7	10.7	15.9	9.3	9.8	39.8
Stauffer	Basal	13.9	25.8	7.1	19.0	7.4	8.3	35.6
	Language Experience	20.8	26.5	7.8	12.4	8.0	8.7	35.1

Table 45 Within-projects analysis of variance and covariance on Stanford measures for the Basal vs. Language Experience comparison

		Word Recognition			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
P[a]		A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]
1	Sex	1.10	.03	.07	3.67	1.30	1.83	.06	2.93	2.06	1.98	.29	.63	1.22	.00	.14
	Treatment	1.45	17.46*	14.10*	.79	10.03*	7.00*	2.09	15.38*	12.42*	1.19	.21	.13	.54	11.34*	7.26*
	S × T	.25	.15	.18	.72	.63	.85	.38	.29	.27	.19	.15	.18	.07	.01	.01
2	Sex	.69	.63	.05	5.18†	1.84	3.59	.00	2.76	1.12	2.88	.44	1.20	.83	.38	.00
	Treatment	1.61	2.53	1.78	.21	.14	.00	.63	.72	.84	.45	.49	.16	.54	.65	.33
	S × T	.28	.52	.39	.02	.02	.00	.12	.17	.30	.21	.28	.25	.18	.30	.28
3	Sex	4.61†	.51	1.35	8.22††	3.38	3.41	.24	1.73	.43	5.35†	1.22	1.23	4.01†	.29	.68
	Treatment	.38	19.39*	.08	13.07***	.01	6.34**	.01	20.01*	.98	.47	8.25*	.71	1.32	11.92*	1.13
	S × T	.07	.27	.27	.05	.17	.17	.44	.96	.43	.01	.04	.02	.04	.63	.12
4	Sex	1.46	.00	.18	1.56	.04	.18	.11	1.10	.82	2.25	.28	.49	2.42	.22	.70
	Treatment	6.08**	2.68	5.30**	3.53	.46	1.46	1.24	.00	.94	2.09	.61	2.00	.77	.18	.09
	S × T	.11	.01	.01	.05	.00	.04	.94	.93	.25	.17	.04	.00	.80	.80	.38

[a] Projects in numerical order are Cleland, Hahn, Kendrick, and Stauffer.
[b] Summarizes analysis of variance. *F* ratios are based on 1 and 226 *df*
[c] Summarizes covariance using Phonemes and Identical Forms as covariates. *F* ratios are based on 1 and 224 *df*
[d] Summarizes covariance using all seven premeasures as covariates. *F* ratios are based on 1 and 219 *df*
* Significant difference favoring Language Experience at .01 level
** Significant difference favoring Basal at .05 level
*** Significant difference favoring Basal at .01 level
† Significant difference favoring females at .05 level
†† Significant difference favoring females at .01 level

BEST COPY AVAILABLE

Table 46 Unadjusted and adjusted Stanford means for the Basal vs. Language Experience comparison

		Word Reading			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
Project	Trt	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c
Cleland	Basal	22.6	22.7	22.5	25.5	25.7	25.7	23.0	23.1	23.1	16.0	16.5	15.8	39.1	39.3	39.1
	LE	24.1	25.5	25.5	26.8	28.8	28.8	24.6	25.8	25.7	14.8	15.6	15.5	40.4	42.4	42.2
Hahn	Basal	22.0	22.7	21.8	21.8	21.9	21.0	21.7	21.8	20.6	13.2	13.5	12.8	38.5	38.9	39.5
	LE	24.0	23.8	22.8	22.5	22.1	21.1	22.5	22.3	21.2	13.9	13.9	13.1	39.8	39.6	38.1
Kendrick	Basal	19.3	16.6	17.8	19.7	16.3	17.6	21.5	19.3	21.2	10.2	8.5	9.4	34.8	31.2	33.1
	LE	18.8	18.0	18.0	15.9	15.1	15.6	21.6	21.0	20.6	9.6	8.9	8.8	33.5	32.4	32.2
Stauffer	Basal	16.0	20.9	20.2	15.9	21.9	20.8	18.4	22.5	21.2	9.0	12.7	11.6	33.8	40.7	39.1
	LE	19.2	21.7	22.2	19.1	22.0	22.4	19.7	21.7	22.0	10.7	12.7	13.2	35.4	38.9	39.4

[a] Reports unadjusted means

[b] Means adjusted for premeasure difference on Phonemes and Learning Rate

[c] Means adjusted for all seven covariates

Table 47 Subjects used for the analysis of individual outcome measures for the Basal vs. Language Experience comparisons

Project	Treatment	Males	Females	Total
Hahn	Basal	24	26	50
	Language Experience	31	20	51
Kendrick	Basal	25	24	49
	Language Experience	24	25	49
Stauffer	Basal	20	19	39
	Language Experience	19	15	34

However, the extent of the superiority varied from project to project, thereby bringing about a significant project by treatment interaction effect. On the Paragraph Meaning subtest, the Language Experience treatment was superior in three of four projects although the extent of the superiority varied considerably. However, in the fourth project, the Basal treatment was superior. A similar situation occurred with respect to the Vocabulary subtest. On the Word Study Skills subtest, the difference in two projects favored the Basal treatment, but, in the other two projects, the difference favored the Language Experience group. Therefore, on all of the subtests except the Word Reading test, the significant project by treatment interaction was caused by treatment differences favoring one approach in one project, the other approach in another project.

Because of the treatment by project interactions, it was necessary to perform a within-project analysis. The analysis of variance on premeasures is reported in Table 43. A number of treatment differences were found in the various projects. Evidently, the process of assigning students or classes to treatments did not achieve the desired result of placing pupils of equal prereading capability in the two treatment groups. Relatively few sex differences were found to be significant in these four projects. The actual mean performance of the various treatment groups within projects on the readiness measures is reported in Table 44.

The analysis of variance and covariance on the Stanford measures is reported in Table 45. One striking finding again is absence of significant sex by treatment interactions in any of the projects. Neither the Basal nor Language Experience approach was uniquely advantageous or disadvantageous for boys or girls. Treatment differences generally favored the Language Experience approach. Columns C of Table 45 report the covariance analysis using all seven premeasures as covariates. In this analysis, two significant differences favoring the Language Experience approach were found for the Word Reading test. Two significant differences were likewise found for the Paragraph Meaning subtest, but one of the significant differences favored each of the two treatments. No significant differences were found on the Spelling test, and one of the four projects found a significant difference favoring the Language Experience treatment on both the Vocabulary and Word Study Skills tests.

The unadjusted and adjusted means for each of the analyses are reported in Table 46. The adjusted means are generally quite similar for treatment groups in each of the projects. It is unlikely that even those differences which were found to be statistically significant were of much practical significance.

Analysis of individual outcome measures. The projects which were used to analyze Language Experience versus Basal treatment differences on the individual tests are listed in Table 47. This table records the number of students who comprised the sample for each treatment within each project. One of the four projects which was used in the analysis of Stanford measures is not included in this analysis because of the unavailability of sample data. An indication of the comparability of the Language Experience and Basal subjects within projects can be ascertained from Table 48, which summarizes an analysis of variance conducted on premeasures. In two of the three projects, significant treatment differences on certain premeasures are indicated.

The analysis of variance and covariance on the individual outcome measures is presented in Table 49. Again, there is no evidence of sex by treatment interaction in the second covariance analysis. Furthermore, in

Table 48 Within-projects analysis of variance on premeasures for Basal vs. Language Experience comparison

Project	Effect	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Pattern Copying	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Hahn	Sex	.00	2.63	1.16	.23	1.97	.82	.56	.02
	Treatment	1.78	.01	.18	.09	2.18	.29	.02	.60
	S × T	.12	.10	.30	1.67	.74	1.49	.09	.12
Kendrick	Sex	2.32	.30	2.66	.36	9.77††	.01	3.26	.05
	Treatment	5.64***	1.74	1.38	.05	.80	4.13*	4.06*	5.46*
	S × T	.04	.12	.28	.10	.12	.51	.18	.00
Stauffer	Sex	14.89††	14.59††	2.21	1.95	.92	.03	7.26††	12.75††
	Treatment	7.60**	.70	1.59	1.87	9.92†	2.96	.27	.06
	S × T	.07	.40	1.03	.05	.09	.11	1.21	.00

* Significant difference favoring Language Experience at .05 level.
** Significant difference favoring Language Experience at .01 level.
*** Significant difference favoring Basal at .05 level.
† Significant difference favoring Basal at .01 level.
†† Significant difference favoring females at .01 level.
Note: All *F* ratios are based on 1 and 260 *df*.

Table 49 Within-projects analysis of variance and covariance on individual outcome measures for Basal vs. Language Experience comparison

		Gilmore Accuracy			Gilmore Rate			Fry Word List			Gates Word List		
P[a]	Effect	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]
1	Sex × Treatment	.94	2.88	2.49	1.14	1.80	2.09	1.52	4.10††	3.47	1.40	4.14††	3.47
	Treatment	.56	.01	.01	.02	.06	.10	.14	1.98	1.76	.18	.09	.04
	Sex	3.33	2.02	2.30	6.37**	4.79**	4.71**	.36	.00	.01	.93	.15	.23
2	Sex × Treatment	.31	.13	.11	1.03	.87	.67	.04	.33	.30	.05	.00	.00
	Treatment	1.90	2.79	3.30	.10	.34	.43	.01	.00	.02	.01	.23	.12
	Sex	4.62**	1.66	1.63	1.54	.53	.50	1.42	.10	.06	1.64	.21	.11
3	Sex × Treatment	.46	.19	.20	.01	.12	.13	1.56	1.49	1.42	1.21	1.01	.88
	Treatment	10.17*	14.74*	13.14*	.76	1.08	.94	25.02*	33.57*	30.81*	13.01*	20.12*	17.36*
	Sex	8.40***	.32	.53	3.19	.00	.05	1.65	4.48†	3.93†	4.58**	.83	.46

[a] Projects in numerical order are Hahn, Kendrick, and Stauffer.
[b] Summarizes analysis of variance. *F* ratios are based on 1 and 260 *df*.
[c] Summarizes covariance using Phonemes, Letter Names, Identical Forms, Meaning, and Listening as covariates. *F* ratios are based on 1 and 255 *df*.
[d] Summarizes covariance using all eight premeasures as covariates. *F* ratios are based on 1 and 252 *df*.
* Significant difference favoring Language Experience at .01 level.
** Significant difference favoring females at .05 level.
*** Significant difference favoring females at .01 level.
† Significant difference favoring males at .01 level.
†† Project difference or interaction significant at .05 level.

BEST COPY AVAILABLE

Table 50 Unadjusted and adjusted means on individual outcome measures for the Basal vs. Language Experience comparison

		Gilmore Accuracy			Gilmore Rate			Fry Word List			Gates Word List		
Project	Treatment	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]
Hahn	Basal	24.9	23.5	24.1	67.6	66.0	65.4	10.4	9.4	10.0	15.6	14.6	15.3
	LE	23.4	23.4	23.9	67.1	66.9	66.7	11.0	11.0	11.5	15.0	14.9	15.5
Kendrick	Basal	14.6	12.7	11.9	35.3	33.3	34.2	4.9	3.0	2.3	10.6	9.3	8.5
	LE	17.4	15.6	15.1	34.0	30.9	31.5	4.8	3.1	2.5	10.4	8.8	8.1
Stauffer	Basal	17.1	20.4	20.7	53.7	58.0	57.7	2.5	6.0	6.2	10.1	12.9	13.2
	LE	24.6	28.1	28.1	57.9	62.9	62.3	11.5	14.3	14.2	16.0	18.7	18.6

[a] Reports unadjusted means

[b] Means adjusted for Phonemes, Letter Names, Identical Forms, Meaning, and IQ

[c] Means adjusted for all eight premeasures

Table 51 Materials and numbers of classes and pupils for Basal vs. Linguistic

	Ruddell		Schneyer		Sheldon	
	Classes	Pupils	Classes	Pupils	Classes	Pupils
Numbers						
Basal	6	120	12	334	7	143
Linguistic	5	97	12	347	14	316
Materials						
Basal	Allyn-Bacon		Scott-Foresman		Ginn	
Linguistic	McGraw-Hill		Fries		Singer Bloomfield-Barnhart	

most cases, sex differences were not found in rate of reading and only one difference, that favoring the Language Experience approach, was found in Reading Accuracy. One of the three projects reported a significant difference favoring the Language Experience approach on both the Fry and Gates word lists. The unadjusted and adjusted means for each treatment group within each project are recorded in Table 50. It is difficult to ascertain any definite trend regarding the effectiveness of the two treatments in terms of achievement on the individual measures.

Summary of Basal versus Language Experience comparison. Relatively few significant differences were found between the Language Experience and Basal approaches. Those significant differences which were found to exist generally favored the Language Experience approach. However, these sporadic differences were often not of much practical significance in terms of actual reading achievement. Little was found in this analysis to support a claim of superiority by either the Language Experience or Basal method.

Basal versus Linguistic comparisons

Three projects were involved in the assessment of the relative effectiveness of Basal and Linguistic programs. Information concerning the number of students and types of materials which comprised the Basal and Linguistic groups is provided in Table 51. Three different sets of basal readers were used in the three projects and four types of linguistic readers were employed in the Linguistic group. Again, the assumption was made that the Basal programs had a great deal in common with one another and that the Linguistic programs also had many similarities.

The across-projects analysis of variance on the premeasures is reported in Table 52. This table is surprisingly devoid of significant effects. Only the main effect for projects is significant, thereby indicating that pupils in the various projects differed considerably in their readiness for reading. The across-projects analysis of variance and covariance on Stanford measures is reported in Table 53. Here again, the project differences are most striking even when pupil readiness is adjusted by covari-

Table 52 Across-projects analysis of variance on premeasures for Basal vs. Linguistic comparison

Effect	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Sex	.56	.78	.13	1.47	1.13	.14	.16
Treatment	2.13	.99	3.24	.78	1.23	2.36	.76
Sex × Treatment	.00	.01	.08	.23	.09	.00	.00
Project	19.73**	3.99**	6.56**	2.26	16.66**	16.23**	17.70**
Sex × Project	.15	.50	.19	.11	.13	.26	.39
Trt. × Project	.26	.10	.88	2.30	.99	.29	.21
Sex × Trt. × Proj.	.03	.17	.16	.12	.16	.12	.00

* Project difference or interaction significant at .05 level.
** Project difference or interaction significant at .01 level.
Note: Numerator *df* = 1 except for all effects involving project in which case numerator *df* = 2. Denominator *df* = 100.

Table 53 Across-projects analysis of variance and covariance on Stanford measures for Basal vs. Linguistic comparison

	Word Recognition			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
Effect	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]
Sex	.09	.00	.56	.57	.55	1.70	.04	.49	.00	1.34	1.81	2.73	.94	1.17	2.96
Treatment	.10	3.36	4.38*	4.09**	.96	6.12**	1.69	.00	.04	2.50	.11	.49	1.74	.01	.22
S × T	.01	.29	.34	.10	.71	.82	.07	.51	.74	.06	.00	.09	.05	.47	.46
Project	3.18†	1.79	12.71††	1.60	4.94†	15.55††	8.89††	12.15††	2.11	3.27†	8.21††	9.22††	5.77††	10.15††	8.30††
S × P	.04	.34	.50	.06	.30	.39	.02	.09	.34	.06	.06	.73	.10	.09	.27
T × P	1.08	8.64††	11.32††	.50	3.14	.97	.23	2.75	1.85	1.25	7.42††	5.75††	.55	5.04††	4.72††
S × T × P	.08	.32	.82	.16	.28	1.06	.12	.31	.66	.01	.14	.12	.01	.16	.22

[a] Summarizes analysis of variance. Numerator *df* = 1 except for those involving project which are based on 2 *df*. Denominator *df* = 100.
[b] Summarizes covariance using Learning Rate as covariate. Numerator *df* = 1 except for those involving project which are based on 2 *df*. Denominator *df* = 99.
[c] Summarizes covariance using all seven premeasures as covariances. Numerator *df* = 1 except for those involving project which are based on 2 *df*. Denominator *df* = 93.
* Significant difference favoring Linguistic at .05 level.
** Significant difference favoring Basal at .05 level.
† Project difference or interaction significant at .05 level.
†† Project difference or interaction significant at .01 level.

Table 54 Within-projects analysis of variance on premeasures for the Basal vs. Linguistic comparison

Project	Effect	Murphy-Durrell Phonemes	Murphy-Durrell Total Letters	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
	Sex × Treatment	.02	.05	.06	.35	.43	.22	.01
Ruddell	Treatment	.48	.07	4.38*	1.15	2.56	2.72	1.21
	Sex	.03	.00	.05	.09	.30	.05	.31
	Sex × Treatment	.02	.03	.38	.04	.00	.02	.01
Schneyer	Treatment	2.14	1.17	.60	2.07	.12	1.26	.47
	Sex	.44	.01	.03	.30	.94	.00	.10
	Sex × Treatment	.02	.28	.02	.00	.04	.00	.00
Sheldon	Treatment	4.06*	1.00	1.05	1.81	3.56	2.39	1.81
	Sex	.40	1.76	.37	1.38	.09	.60	.52

* Significant difference favoring Basal at .05 level.
Note: All *F* ratios are based on 1 and 100 *df*.

BEST COPY AVAILABLE

Table 55 Premeasure means for the Basal vs. Linguistic comparison

Project	Treatment	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Ruddell	Basal	22.2	31.4	11.5	14.3	8.6	9.1	36.1
	Linguistic	19.3	30.2	9.5	17.0	7.1	8.0	32.2
Schneyer	Basal	19.7	30.2	9.1	12.1	7.5	8.1	30.1
	Linguistic	15.6	27.0	8.6	14.6	7.3	7.6	28.5
Sheldon	Basal	34.0	36.7	10.7	17.7	10.9	10.3	41.9
	Linguistic	27.6	33.2	10.0	15.1	9.5	9.4	38.2

Table 56 Within-project analysis of variance and covariance on Stanford measures for the Basal vs. Linguistic comparison

		Word Recognition			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
P^a		A^b	B^c	C^d	A^b	B^c	C^d	A^b	B^c	C^d	A^b	B^c	C^d	A^b	B^c	C^d
1	Sex	.01	.24	.36	.00	.06	.61	.01	.17	.22	.49	.63	3.44	.05	.01	1.76
	Treatment	.98	18.69**	25.37**	.40	2.26	.34	.12	3.05	2.54	.25	9.55**	8.42**	.01	6.61*	6.48*
	S × T	.11	.04	2.07	.34	.37	3.08	.00	.08	.38	.00	.08	.14	.03	.00	.82
2	Sex	.04	.35	.97	.23	.95	1.67	.08	.04	.01	.27	1.00	.70	.18	.78	1.28
	Treatment	1.24	.67	.71	1.22	.62	1.04	1.96	1.49	1.14	3.26	3.41	2.61	2.63	2.53	2.92
	S × T	.01	.87	.10	.02	.91	.12	.01	.25	.03	.04	.16	.07	.01	.73	.02
3	Sex	.12	.05		.48	.12		.00	.38		.68	.30		.93	.59	
	Treatment	.38	.12	.04	4.68‡	4.54‡	.08	2.47	1.43	.18	3.05	2.14	.24	2.09	1.03	.58
	S × T	.06	.06	.04	.03	.01	.00	.29	.88	1.88	.05	.03	.05	.00	.03	.20

[a] Projects in numerical order are Ruddell, Schneyer, and Sheldon

[b] Summarizes analysis of variance. *F* ratios are based on 1 and 100 *df*

[c] Summarizes covariance using Learning Rate as covariate. *F* ratios are based on 1 and 99 *df*

[d] Summarizes covariance using all seven premeasures as covariates. *F* ratios are based on 1 and 93 *df*

* Significant difference favoring Linguistic at .05 level

** Significant difference favoring Linguistic at .01 level

‡ Significant difference favoring Basal at .05 level

Table 57 Unadjusted and adjusted Stanford means for the Basal vs. Linguistic comparison

		Word Reading			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
Project	Trt.	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c	A^a	B^b	C^c
Ruddell	Basal	18.0	14.2	17.3	17.7	13.0	16.9	19.5	16.5	18.8	8.4	5.4	7.4	32.8	27.8	31.6
	Linguistic	20.6	21.1	22.6	15.6	16.3	17.8	18.7	19.1	20.8	9.5	9.9	11.1	33.1	33.8	35.8
Schneyer	Basal	18.2	19.5	20.2	17.6	19.2	20.3	20.1	21.1	21.7	10.7	11.7	12.3	35.2	37.0	37.9
	Linguistic	16.3	18.6	19.7	15.1	18.1	19.3	18.0	19.9	20.8	8.0	9.9	11.0	31.3	34.5	36.1
Sheldon	Basal	21.2	19.0	15.7	22.3	19.6	15.5	24.8	23.1	20.3	13.3	11.6	9.1	41.0	38.0	33.8
	Linguistic	20.0	19.4	17.6	16.8	16.1	13.8	22.1	21.7	20.2	10.4	10.0	8.5	37.0	36.2	33.6

[a] Reports unadjusted means.

[b] Means adjusted for premeasure differences on Learning Rate

[c] Adjusted for all seven covariates

BEST COPY AVAILABLE

ance. However, treatment differences were found to be significant on the Word Recognition and Paragraph Meaning subtests, the first difference favoring the Linguistic subjects and the second difference favoring the Basal subjects. Treatment by project interactions were found to be significant on three of the five Stanford Achievement measures in the covariance analysis using all seven premeasures as covariates. The explanation for these interactions can be found in Columns C of Table 57. On each of these Stanford subtests for which significant treatment by project interactions were found, the difference favored one treatment in one project and another treatment in another project. The Linguistic and Basal treatments did not operate in the same fashion from project to project.

As a result of the project by treatment interactions, a within-projects analysis was conducted. The analysis of variance on premeasures is reported in Table 54. Only two significant effects are found in the entire table. Table 55 reports the premeasure means for treatment groups within projects. The treatment groups are very similar in performance on the premeasures within each of the projects.

A within-projects analysis of variance and covariance on Stanford measures is reported in Table 56. Again, there are no significant sex by treatment interactions. Neither the Linguistic materials nor Basal materials utilized in these projects has a unique effect on boys and girls. Columns C record the result of a covariance analysis using all seven premeasures as covariates. Relatively few treatment differences were found. One difference favoring the Linguistic approach was found for the Word Reading, Spelling, and Word Study Skills subtests. This general lack of superiority of either approach is further supported by the unadjusted and adjusted means recorded in Table 57. The differences tend to favor one experimental group in one project and the other experimental group in another project.

Table 58 Subjects used for the analysis of individual outcome measures for the Basal vs. Linguistic comparison

Project	Treatment	Males	Females	Total
Ruddell	Basal	21	20	41
	Linguistic	20	14	34
Schneyer	Basal	21	23	44
	Linguistic	16	28	44
Sheldon	Basal	15	20	35
	Linguistic	29	39	68

Analysis of individual outcome measures. The numbers of subjects who comprised the sample group in the three Basal versus Linguistic comparisons are reported in Table 58. The analysis of variance on premeasures for this sample group is reported in Table 59. A number of treatment differences are reported pointing out the difficulty of obtaining experimental groups equal in prereading capability. Surprisingly, two of the three significant sex differences on premeasures favored boys.

Table 59 Within-projects analysis of variance on premeasures for Basal vs. Linguistic comparison

Project	Effect	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Pattern Copying	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Ruddell	Sex	.03	.26	.80	.87	4.65††	6.85†††	.66	.47
	Treatment	.61	.74	3.76	.56	.00	2.62	12.94***	6.65**
	S × T	.96	.01	.34	.08	.37	.38	1.53	.76
Schneyer	Sex	3.42	8.94†	.19	1.55	.20	.03	.00	1.88
	Treatment	1.88	.66	2.12	3.95**	6.20*	.32	1.41	.35
	S × T	.12	.12	2.12	.45	.87	.18	.02	2.06
Sheldon	Sex	.96	.07	.11	.09	1.91	.03	1.23	2.18
	Treatment	5.13**	4.27**	.40	.13	3.11	4.61**	1.78	4.72**
	S × T	1.59	.20	.09	.76	1.02	1.69	1.42	1.83

* Significant difference favoring Linguistic at .05 level
** Significant difference favoring Basal at .05 level
*** Significant difference favoring Basal at .01 level
† Significant difference favoring females at .01 level
†† Significant difference favoring males at .05 level
††† Significant difference favoring males at .01 level
Note. All *F* ratios are based on 1 and 254 *df*

BEST COPY AVAILABLE

Table 60 Within-projects analysis of variance and covariance on individual outcome measures for Basal vs. Linguistic comparison

P[a]	Effect	Gilmore Accuracy A[b]	Gilmore Accuracy B[c]	Gilmore Accuracy C[d]	Gilmore Rate A[b]	Gilmore Rate B[c]	Gilmore Rate C[d]	Fry Word List A[b]	Fry Word List B[c]	Fry Word List C[d]	Gates Word List A[b]	Gates Word List B[c]	Gates Word List C[d]
1	Sex	.02	.94	1.04	.12	1.92	1.46	.45	.02	.01	.98	.01	.11
	Treatment	1.98	.02	.06	2.74	.20	.56	.99	8.90**	6.26*	.36	1.32	.31
	Sex × Treatment	.07	.02	.07	.01	.00	.03	.01	.07	.60	.41	1.11	2.94
2	Sex	8.46†††	9.49†††	5.86††	7.59†††	7.53†††	3.55	5.43††	5.10††	1.94	5.33††	5.26††	1.72
	Treatment	3.46	1.34	2.32	5.34***	3.17	4.47***	.46	.41	.25	2.35	.41	.85
	Sex × Treatment	.04	.33	.36	.77	1.83	2.62	.04	.22	.11	.25	.83	.94
3	Sex	1.03	.34	.41	1.11	.45	.48	.00	.36	.36	.05	.20	.27
	Treatment	7.14†	6.08***	2.03	9.68†	8.57†	4.67***	3.62	10.05**	24.17**	.37	.03	2.85
	Sex × Treatment	.00	.79	1.65	.16	.06	.14	.01	.81	1.81	.05	1.61	3.18

[a] Projects in numerical order are Ruddell, Schneyer, and Sheldon.
[b] Summarizes analysis of variance. *F* ratios are based on 1 and 254 *df*.
[c] Summarizes covariance using Learning Rate, Pattern Copying, Identical Forms, and Listening as covariates. *F* ratios are based on 1 and 250 *df*.
[d] Summarizes covariance using all eight premeasures as covariates. *F* ratios are based on 1 and 246 *df*.
* Significant difference favoring Linguistic at .05 level.
** Significant difference favoring Linguistic at .01 level.
*** Significant difference favoring Basal at .05 level.
† Significant difference favoring Basal at .01 level.
†† Significant difference favoring females at .05 level.
††† Significant difference favoring females at .01 level.

Table 61 Unadjusted and adjusted means on individual outcome measures for the Basal vs. Linguistic comparison

Project	Trt.	Gilmore Accuracy A[a]	Gilmore Accuracy B[b]	Gilmore Accuracy C[c]	Gilmore Rate A[a]	Gilmore Rate B[b]	Gilmore Rate C[c]	Fry Word List A[a]	Fry Word List B[b]	Fry Word List C[c]	Gates Word List A[a]	Gates Word List B[b]	Gates Word List C[c]
Ruddell	Basal	19.1	15.5	17.8	54.0	46.3	50.7	6.2	3.9	5.6	11.5	9.1	10.9
	Linguistic	15.0	15.8	17.3	42.4	43.6	46.3	8.0	8.1	8.9	10.4	10.7	11.6
Schneyer	Basal	24.2	25.0	27.0	60.5	62.4	66.1	6.8	6.9	8.1	11.6	12.0	13.3
	Linguistic	19.1	22.5	24.0	45.7	52.6	55.1	5.7	7.7	8.7	8.9	11.1	12.2
Sheldon	Basal	26.7	24.8	20.3	62.9	59.1	50.8	6.6	5.9	3.0	13.1	12.0	8.8
	Linguistic	19.6	19.7	17.5	43.4	43.7	40.0	9.6	9.9	8.5	12.1	12.3	10.7

[a] Reports unadjusted means.
[b] Means adjusted for Learning Rate, Pattern Copying, Identical Forms, and Listening.
[c] Means adjusted for all eight premeasures.

The within-projects analysis of variance and covariance on individual outcome measures is recorded in Table 60. Columns C of the table summarize an analysis of covariance which utilized all eight premeasures as covariates. In this analysis, no treatment differences were found in reading accuracy. Two of the three projects found significant differences favoring the Basal group in rate of reading. Conversely, two significant differences favoring the Linguistic group were found in performance on the Fry Phonetically Regular Word Test. This finding is somewhat to be expected because the Fry list was devised to approximate the vocabulary introduced in Linguistic programs. It was thought the Gates list would favor pupils who had been taught to read using a Basal series. However, in this analysis no treatment differences were found on the Gates test. The unadjusted and adjusted means for the Linguistic and Basal groups in the three projects are listed in Table 61. The means indicate that the Basal method produced higher mean performance in terms of reading accuracy in all three projects although these differences were not significant. Likewise, the Basal program produced higher rate of reading in all three projects, two cases of which proved to be statistically significant. The Linguistic group outperformed the

BEST COPY AVAILABLE

Basal group on the Fry test in all three projects. The Linguistic group also surpassed the Basal group on the Gates test in two of the three projects.

Summary of Basal versus Linguistic comparison. The most common finding for the Linguistic versus Basal comparison was that of no difference between treatments. However, the Linguistic group tended to outperform the Basal group on tests of word recognition while the Basal group exhibited somewhat greater speed and accuracy in reading. No differences in comprehension were ascertained.

Basal versus Phonic/Linguistic comparisons

Three projects were involved in the Basal versus Phonic/Linguistic comparison. Information about the numbers of classes and students and the kinds of basal series utilized is recorded in Table 62. Two of the three projects used the same basal series, but the third project used a variety of basal materials. The Phonic/Linguistic series was treated as a separate approach because it did not seem to fit any of the other categories used in this investigation.

The across-projects analysis of variance on premeasures is reported in Table 63. Significant sex differences favoring girls were found on five of the seven premeasures. Significant treatment effects favoring the Phonic/Linguistic subjects were found on two of the seven premeasures. Highly reliable project differences were reported. Also, treatment by project interactions were found to be significant on three of the seven premeasures.

The analysis of variance and covariance on Stanford measures across projects is reported in Table 64. Although the analysis of variance summarized in Columns A found significant sex differences favoring females on four of the five outcome measures, these differences were erased when the achievement scores were adjusted for differences in prereading capability. In other words, the superiority of girls in reading capability at the end of the year appeared to be merely a reflection of their superiority in readiness for reading at the beginning of the year. Project differences were found to be highly reliable, indicating that pupils differed considerably from project to project in their reading ability. Again, sex by treatment interactions were found to be negligible.

Table 62 Materials and numbers of classes and pupils for Basal vs. Phonic/Linguistic

	Hayes		Tanyzer		Wyatt	
	Classes	Pupils	Classes	Pupils	Classes	Pupils
Numbers						
Basal	5	87	9	228	10	210
Phonic/Linguistic	5	99	8	186	10	203
Materials						
Basal	Scott-Foresman		Scott-Foresman		Variety	

Table 63 Across-projects analysis of variance on premeasures for Basal vs. Phonic/Linguistic comparison

Effect	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Sex	7.28***	8.34***	5.10**	5.40**	3.06	3.62	8.69***
Treatment	15.04*	.57	11.28*	2.31	2.27	1.24	3.39
Sex × Treatment	.07	.08	1.32	.23	.13	.54	.24
Project	51.31††	51.53††	42.41††	10.90††	34.45††	34.15††	31.23††
Sex × Project	.13	.95	1.73	.11	.66	.46	.93
Trt. × Project	8.45††	3.03	.39	6.85††	3.82†	2.38	2.28
Sex × Trt. × Proj.	.55	.38	.11	.27	.72	.85	.33

* Significant difference favoring Phonic/Linguistic at .01 level
** Significant difference favoring females at .05 level
*** Significant difference favoring females at .01 level
† Project difference or interaction significant at .05 level
†† Project difference or interaction significant at .01 level
Note. Numerator *df* = 1 except for all effects involving project in which case numerator *df* = 2. Denominator *df* = 82.

BEST COPY AVAILABLE

Table 64 Across-projects analysis of variance and covariance on Stanford Measures for Basal vs. Phonic/Linguistic comparison

	Word Recognition			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
Effect	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]
Sex	7.60†	.95	.39	13.60†	4.46***	1.41	.05	7.19†††	6.35††	8.00†	1.27	.01	4.90***	.00	.19
Treatment	77.20**	54.17**	79.89**	18.18**	4.24*	8.95**	5.72*	.07	1.45	20.16**	5.39*	10.59**	19.33**	3.81	7.01**
S × T	.03	.00	.17	.45	.00	.29	.14	.04	.29	.01	.03	.02	.05	.49	.20
Project	6.71*†	14.70*†	24.15*†	6.21*†	18.87*†	32.58*†	4.26†*	11.97*†	23.07*†	6.13*†	6.59*†	12.59*†	5.56*†	17.00*†	20.72*†
S × P	.44	.49	.36	.24	.25	.60	.64	.99	1.10	.64	.70	.18	.30	.15	.13
T × P	3.99†*	1.57	3.32†*	2.64	1.12	1.51	3.40†*	.15	.36	1.32	2.28	1.16	3.09	2.05	1.69
S × T × P	.31	.02	.32	.04	.11	.30	.36	.40	.08	.65	.28	.36	.22	.01	.08

[a] Summarizes analysis of variance. Numerator *df* = 1 for all effects except those involving project based on 2 *df*. Denominator *df* = 2.

[b] Summarizes covariance using Phonemes and Identical Forms as covariates. Numerator *df* = 1 for all effects except those involving project based on 2 *df*. Denominator *df* = 80.

[c] Summarizes covariance using all seven premeasures as covariates. Numerator *df* = 1 for all effects except those involving project based on 2 *df*. Denominator *df* = 75.

* Significant difference favoring Phonic/Linguistic at .05 level.

** Significant difference favoring Phonic/Linguistic at .01 level.

*** Significant difference favoring females at .05 level.

† Significant difference favoring females at .01 level.

†† Significant difference favoring males at .05 level.

††† Significant difference favoring males at .01 level.

†* Project difference or interaction significant at .05 level.

*† Project difference or interaction significant at .01 level.

Table 65 Within-projects analysis of variance on premeasures for the Basal vs. Phonic/Linguistic comparison

Project	Effect	Murphy-Durrell Phonemes	Murphy-Durrell Total Letters	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
	Sex × Treatment	.65	.21	.43	.03	.01	.59	.00
Hayes	Treatment	.02	1.14	1.58	2.62	2.00	.24	.28
	Sex	1.41	3.81	3.23	1.46	.22	1.12	3.39
	Sex × Treatment	.20	.62	.11	.71	.26	1.26	.85
Tanyzer	Treatment	32.10**	5.70*	8.00**	13.36**	7.25**	.10	7.21**
	Sex	4.08†	5.76†	5.26†	2.83	.14	3.17	6.41†
	Sex × Treatment	.31	.00	1.08	.03	1.31	.37	.03
Wyatt	Treatment	.62	.01	3.77	.01	.89	5.75*	.60
	Sex	2.06	.70	.01	1.34	4.01†	.28	.78

* Significant difference favoring Phonic/Linguistic at .05 level.

** Significant difference favoring Phonic/Linguistic at .01 level.

† Significant difference favoring females at .05 level.

Note: All *F* ratios are based on 1 and 82 *df*.

Table 66 Premeasure means for the Basal vs. Phonic/Linguistic comparison

Project	Trt.	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Hayes	Basal	14.7	24.0	7.0	13.1	7.2	7.8	33.3
	Lipp.	15.0	21.1	7.9	10.6	6.4	7.6	32.4
Tanyzer	Basal	24.0	36.7	10.9	12.4	8.8	9.7	37.6
	Lipp.	35.4	41.8	12.5	16.8	10.0	9.6	41.2
Wyatt	Basal	29.3	37.6	9.8	16.1	9.4	9.7	40.6
	Lipp.	30.8	37.5	10.8	16.2	9.8	10.5	41.6

BEST COPY AVAILABLE

Treatment differences were found to favor the Phonic/Linguistic approach on four of the five Stanford measures. Moreover, with respect to the covariance analysis, only one treatment by project interaction was found to be significant. Apparently the Phonic/Linguistic and Basal treatments operated in a similar fashion from project to project. Furthermore, the Phonic/Linguistic treatment tended to produce higher reading achievement at the end of the first grade.

Despite the relative freedom from project by treatment interaction, a within-projects analysis was conducted. This analysis was performed in the interests of consistency with the other Basal versus nonbasal treatment comparisons. The within-projects analysis of variance on premeasures is reported in Table 65. Two of the three projects are free from significant treatment effects. However, the other project found significant treatment differences favoring the Phonic/Linguistic group on six of the seven premeasures. Obviously, in the project, the Phonic/Linguistic group was in a very favored position in terms of readiness for reading. Further information concerning this is presented in Table 66 which reports the premeasure means for treatments within projects. In the project in question, large differences were found in mean performance between treatments on most of the premeasures. This lack of homogeneity between treat-

Table 67 Within-projects analysis of variance and covariance on Stanford measures for the Basal vs. Phonic/Linguistic comparison

		Word Recognition			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
P[a]		A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]
1	Sex	3.58	1.66	.32	4.78†	2.70	.37	.69	.01	.50	4.15†	2.17	.06	1.29	.04	.32
	Treatment	15.27**	26.44**	32.27**	2.98	5.38*	5.27*	.06	.03	.07	5.17*	8.74**	8.34**	3.21	7.50**	7.54**
	S × T	.38	.03	.11	.02	.14	.18	.06	.10	.00	.76	.27	.43	.49	.05	.13
2	Sex	3.60	.30	.02	5.57†	1.32	.03	.03	4.21†	7.52††	4.04†	.56	.22	3.42	.08	.17
	Treatment	54.16**	13.03**	31.17**	18.21**	.22	4.16*	11.92**	.32	1.38	14.38**	.00	2.03	20.41**	.00	1.58
	S × T	.18	.00	.00	.02	.08	.07	.80	.45	.39	.53	.21	.27	.02	.17	.23
3	Sex	1.30	.02	.79	3.73	1.32	2.56	.61	5.41†	2.40	1.09	.00	.01	.78	.16	.00
	Treatment	14.67**	18.59**	17.49**	1.58	.97	.39	.37	.02	.03	3.05	2.54	2.24	1.58	1.00	.55
	S × T	.09	.00	.71	.09	.00	.64	.01	.28	.10	.00	.11	.05	.00	.30	.00

[a] Projects in numerical order are Hayes, Tanyzer, and Wyatt
[b] Summarizes analysis of variance. *F* ratios are based on 1 and 82 *df*
[c] Summarizes covariance using Phonemes and Identical Forms as covariates. *F* ratios are based on 1 and 82 *df*
[d] Summarizes covariance using all seven premeasures as covariates. *F* ratios are based on 1 and 75 *df*
* Significant difference favoring Phonic/Linguistic at .05 level
** Significant difference favoring Phonic/Linguistic at .01 level
† Significant difference favoring females at .05 level
†† Significant difference favoring females at .01 level

Table 68 Unadjusted and adjusted Stanford means for the Basal vs. Phonic/Linguistic comparison

		Word Recognition			Paragraph Meaning			Vocabulary			Spelling			Word Study Skills		
Project	Trt	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]
Hayes	Basal	18.5	23.6	24.7	19.3	26.1	28.0	22.0	26.6	27.3	9.0	13.4	14.3	34.2	42.7	48.9
	P/L	25.2	30.4	31.4	23.5	30.6	32.0	21.6	26.4	27.6	12.5	17.0	17.6	38.8	47.7	48.6
Tanyzer	Basal	17.5	18.9	17.2	16.1	18.0	15.1	20.5	21.7	20.3	10.2	11.4	9.8	33.9	36.2	34.3
	P/L	27.2	23.3	23.4	24.1	18.8	18.5	24.7	21.1	21.4	14.7	11.4	11.3	42.8	36.2	36.4
Wyatt	Basal	22.7	21.4	21.8	23.3	21.5	22.4	24.0	22.8	22.8	13.1	11.9	12.4	40.3	38.0	38.4
	P/L	27.4	25.4	25.4	25.5	22.9	23.2	24.7	22.9	22.9	15.0	13.3	13.6	42.5	39.3	39.3

[a] Reports unadjusted means
[b] Means adjusted for premeasure differences on Phonemes and Identical Forms
[c] Means adjusted for all seven covariates

BEST COPY AVAILABLE

ment groups must be considered in interpreting the achievement results.

The within-projects analysis of variance and covariance on the Stanford measures is reported in Table 67. The general superiority of the Phonic/Linguistic program is also indicated here, although the results are not as striking as in the across-projects analysis. In the within-projects analysis, all significant treatment differences favored the Phonic/Linguistic program. In the second covariance analysis, all three projects reported significant differences in favor of the Phonic/Linguistic program on the Word Reading test. Two of the three projects found significant differences favoring the Phonic/Linguistic program on the Paragraph Meaning test. One of the three projects found significant differences favoring the Phonic/Linguistic program on both the Spelling and Word Study Skills tests. No significant differences were found between treatments on the Vocabulary test. A further indication of this same trend can be found in Table 68 which reports the unadjusted and adjusted means for the various projects. Each of the mean comparisons in Columns C for each outcome variable within each project favored the Phonic/Linguistic program. In other words, even those differences which were not found to be statistically significant pointed out the same general trend.

Analysis of individual outcome measures. The number of subjects who were given the individual tests in the Basal versus Phonic/Linguistic projects are recorded in Table 69. The result of the within-projects analysis of variance on the premeasure scores of these individuals is reported in Table 70. Nothing very unusual is found there except that the treatment differences favoring the Phonic/Linguistic program on premeasures in one project are again pointed out.

The within-projects analysis of variance and covariance on the individual outcome measures is reported in Table 71. Columns C report an analysis of covariance using all eight premeasures as covariates. In this analysis, no differences between treatments were found in reading rate, while one difference favoring the Phonic/Linguistic program was found in reading accuracy. However, striking differences in favor of the Phonic/Linguistic program were found in each project for each of the word recognition tests. The corresponding unadjusted and adjusted means are reported in Table 72 and tend to lend further

Table 69 Subjects used for the analysis of individual outcome measurs for the Basal vs. Phonic/Linguistic comparison

Project	Treatment	Males	Females	Total
Hayes	Basal	15	15	30
	Phonic/Linguistic	16	14	30
Tanyzer	Basal	9	8	17
	Phonic/Linguistic	8	6	14
Wyatt	Basal	25	25	50
	Phonic/Linguistic	25	25	50

Table 70 Within-projects analysis of variance on premeasures for Basal vs. Phonic/Linguistic comparison

Project	Eff.	Murphy-Durrell Phonemes	Murphy-Durrell Letter Names	Murphy-Durrell Learning Rate	Thurstone Pattern Copying	Thurstone Identical Forms	Metropolitan Meaning	Metropolitan Listening	Pintner-Cunningham IQ
Hayes	Sex	4.44††	8.93†††	4.02††	3.11	5.08††	.29	3.67	6.05††
	Treatment	.54	4.08***	.10	14.50†	8.90†	.25	1.23	1.24
	S × T	2.08	1.26	2.11	.38	.53	.03	.35	.04
Tanyzer	Sex	5.35††	2.20	.23	2.51	.90	.80	.44	6.80††
	Treatment	10.27**	3.45	7.30**	.44	4.96*	3.26	2.34	5.63*
	S × T	1.92	.24	.02	.39	.12	.48	.01	.12
Wyatt	Sex	.80	.92	.03	1.53	2.51	3.97‡*	.07	.88
	Treatment	.11	.68	1.06	2.25	.26	.26	1.89	.01
	S × T	4.14*†	.25	.48	.00	.01	.04	.36	.10

* Significant difference favoring Phonic/Linguistic at .05 level
** Significant difference favoring Phonic/Linguistic at .01 level
*** Significant difference favoring Basal at .05 level
† Significant difference favoring Basal at .01 level
†† Significant difference favoring females at .05 level
††† Significant difference favoring females at .01 level
‡* Significant difference favoring males at .05 level
*† Interaction significant at .05 level
Note: All *F* ratios are based on 1 and 179 *df*

BEST COPY AVAILABLE

support to the superiority of the Phonic/Linguistic program to the Basal approach on the measures utilized in this investigation. Substantial differences between treatments were found on both the Gates and Fry word lists. Furthermore, each mean difference on the Gilmore Accuracy score favors the Phonic/Linguistic program as do two of the three reading rate scores.

Summary of Basal versus Phonic/Linguistic comparison. The data presented here tend to point out the superiority of the Phonic/Linguistic program to the Basal readers utilized in these projects. The Phonic/Linguistic program produced pupils with superior Word Reading, Paragraph Meaning, Spelling, and Word Study Skills. Phonic/Linguistic pupils were also superior on the Fry Test of Phonetically Regular Words and the Gates Word Recognition Test. No significant differences were found between the Phonic/Linguistic and Basal subjects in rate or accuracy of reading.

The practicality of significant differences

Many significant differences have been reported above for the various basal versus nonbasal comparisons. Differences were regarded as being significant if they reached the .05 level of significance. However, with the large number of comparisons involved, one would suspect a substantial number of differences to reach

Table 71 Within-projects analysis of variance and covariance on individual outcome measures for Basal vs. Phonic/Linguistic comparison

		Gilmore Accuracy			Gilmore Rate			Fry Word List			Gates Word List		
P[a]	Effect	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]	A[b]	B[c]	C[d]
1	Sex	5.48†	.49	.75	21.75††	12.63††	13.31††	6.66†	1.13	1.24	6.24†	1.25	1.37
	Treatment	1.48	7.32**	5.35*	.30	.06	.01	39.32**	66.93**	61.89**	18.12**	31.53**	28.35**
	Sex × Treatment	4.31	2.07	1.93	2.60	1.56	1.34	3.64	2.21	1.87	3.65	1.87	1.61
2	Sex	3.53	.43	.55	2.09	.50	.70	.68	.17	.06	1.43	.01	.04
	Treatment	3.25	.22	.01	8.44**	3.81	3.11	30.93**	22.08**	19.74**	12.46**	6.54*	5.33*
	Sex × Treatment	.49	.01	.19	1.14	.62	.90	.15	.03	.00	.03	.16	.04
3	Sex	2.46	1.07	2.81	1.66	.69	1.45	1.33	.19	.62	2.03	.78	1.58
	Treatment	1.89	6.08*	2.94	.20	.01	.18	27.12**	44.44**	37.13**	9.12**	16.22**	12.04**
	Sex × Treatment	1.14	4.27	2.66	.12	.00	.03	.25	.00	.07	.01	.46	.15

[a] Projects in numerical order are Hayes, Tanyzer, and Wyatt

[b] Summarizes analysis of variance. *F* ratios are based on 1 and 179 *df*

[c] Summarizes covariance using Phonemes, Letter Names, Pattern Copying, and Identical Forms as covariates. *F* ratios are based on 1 and 175 *df*

[d] Summarizes covariance using all eight premeasures as covariates. *F* ratios are based on 1 and 171 *df*

* Significant difference favoring Phonic/Linguistic at .05 level

** Significant difference favoring Phonic/Linguistic at .01 level

† Significant difference favoring females at .05 level

†† Significant difference favoring females at .01 level

Table 72 Unadjusted and adjusted means on individual outcome measures for the Basal vs. Phonic/Linguistic comparison

		Gilmore Accuracy			Gilmore Rate			Fry Word List			Gates Word List		
Project	Treatment	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]	A[a]	B[b]	C[c]
Hayes	Basal	19.2	23.7	24.7	65.2	71.3	72.1	3.8	5.8	6.1	11.4	13.7	14.0
	Lipp	23.1	31.3	31.2	62.0	72.7	72.7	16.8	21.6	21.5	20.5	25.1	25.0
Tanyzer	Basal	23.4	20.5	21.0	45.9	43.6	42.7	3.8	2.7	2.5	9.9	8.4	8.4
	Lipp	31.5	22.2	21.3	69.8	59.1	56.9	19.9	14.8	14.1	20.3	15.4	14.7
Wyatt	Basal	30.5	27.6	28.2	57.5	53.5	54.5	10.3	8.7	9.1	15.5	13.9	14.3
	Lipp	33.9	32.7	31.7	55.4	53.1	52.6	18.6	18.0	17.8	20.5	19.9	19.5

[a] Reports unadjusted means

[b] Means adjusted for Phonemes, Letter Names, Pattern Copying, and Identical Forms

[c] Means adjusted for all eight premeasures

BEST COPY AVAILABLE

Table 73 Grade equivalents for Stanford Achievement Test Primary I Battery, Form W

Word Reading		Paragraph Meaning		Vocabulary		Spelling		Word Study Skills	
No. right	Grade score	No. right	Grade score	No. right	Grade score	No. right	Grade score	No. right	Grade score
1	Below 1.0	1	1.0	1	Below 1.0	1	1.0	1	Below 1.0
2		2	1.1	2		2	1.1	2	
3		3	1.1	3		3	1.3	3	
4		4	1.2	4	1.0	4	1.4	4	
5	1.0	5	1.2	5	1.0	5	1.5	5	
6	1.0	6	1.3	6	1.0	6	1.6	6	
7	1.1	7	1.4	7	1.1	7	1.6	7	
8	1.1	8	1.4	8	1.1	8	1.7	8	
9	1.2	9	1.5	9	1.1	9	1.7	9	
10	1.3	10	1.5	10	1.2	10	1.8	10	
11	1.3	11	1.5	11	1.2	11	1.9	11	1.0
12	1.4	12	1.6	12	1.2	12	2.0	12	1.0
13	1.4	13	1.6	13	1.3	13	2.1	13	1.0
14	1.5	14	1.6	14	1.3	14	2.2	14	1.1
15	1.5	15	1.6	15	1.4	15	2.3	15	1.1
16	1.6	16	1.6	16	1.4	16	2.4	16	1.1
17	1.6	17	1.7	17	1.5	17	2.6	17	1.2
18	1.7	18	1.7	18	1.5	18	2.8	18	1.2
19	1.7	19	1.7	19	1.6	19	3.0	19	1.2
20	1.7	20	1.7	20	1.7	20	3.4	20	1.2
21	1.8	21	1.8	21	1.8			21	1.3
22	1.8	22	1.8	22	1.9			22	1.3
23	1.9	23	1.8	23	2.1			23	1.3
24	1.9	24	1.9	24	2.2			24	1.3
25	2.0	25	1.9	25	2.3			25	1.4
26	2.1	26	2.0	26	2.4			26	1.4
27	2.2	27	2.0	27	2.5			27	1.4
28	2.3	28	2.1	28	2.7			28	1.5
29	2.4	29	2.2	29	2.9			29	1.5
30	2.5	30	2.3	30	3.1			30	1.5
31	2.6	31	2.4	31	3.3			31	1.6
32	2.7	32	2.5	32	3.6			32	1.6
33	2.9	33	2.6	33	4.0			33	1.7
34	3.2	34	2.7	34	4.4			34	1.8
35	3.6	35	2.9	35	4.8			35	1.8
		36	3.1	36	5.2			36	1.9
		37	3.6	37	5.5+			37	1.9
		38	4.0	38				38	2.0
				39				39	2.0
								40	2.1
								41	2.2
								42	2.3
								43	2.4
								44	2.5
								45	2.6
								46	2.7
								47	2.8
								48	3.0
								49	3.2
								50	3.4
								51	3.9
								52	4.8
								53	5.5+
								54	
								55	
								56	

statistical significance on the basis of chance alone. Furthermore, a large sample was employed in this investigation. As a result, a relatively small difference between treatments might be statistically significant. It would be of interest to know how important the statistically reliable differences reported are in a practical sense. In the discussion of the results, unadjusted and adjusted means were given for each treatment comparison. These means were based on raw scores for the various achievement tests. Therefore, it is possible to note the degree of disparity between means for the various basal versus nonbasal comparisons. However, since the achievement measures were standardized tests, normative information is also available. Each of the raw scores can be translated into a grade equivalent score. It is, therefore, possible to judge the practical significance of the differences in terms of whether or not the mean achievement for each group would result in similar grade equivalents. Perhaps, two groups could obtain a grade equivalent score of 1.9, even though a statistically significant difference had been obtained in comparing the achievement means.

Relevant information concerning the grade equivalents for various raw scores on each of the Stanford tests is reported in Table 73. This table reveals, for example, that scores of 18, 19, and 20 on the Stanford Word Reading Test result in the same grade equivalent, 1.7. Therefore, two groups would achieve the same grade equivalent even though one of them averaged 18 correct answers on the Word Reading test, while the other group averaged 20 correct answers. A raw score difference of this magnitude in this study, in many instances, would be regarded as statistically significant. The reader is encouraged to check the raw scores obtained on the various tests for each basal versus nonbasal comparison against this table of norms to obtain some idea of the practical significance of the statistically reliable treatment differences reported.

A word should also be said about the average grade equivalents found for the various treatments. The experimental period was 140 days or approximately 7 months. Therefore, a grade score of 1.7 would be a reasonable expectation of achievement. Furthermore, there is some indication that the norms on the Stanford test are somewhat depressed. That is, the same level of achievement on many other primary reading tests would result in a higher grade equivalent.

Variability within treatments

The discussion to this point has centered around the *mean* achievement of pupils in various reading programs. Another important question is the extent to which any program reduces or increases pupil variability. Two kinds of information from this study are relevant. In the first place, the within-projects analysis made possible the location of the highest and lowest mean class achievement within each treatment. Assuming that classes were randomly assigned to treatments, it might be expected that a superior treatment would tend to be superior across all classrooms. Perhaps, as a result, the classes involved in the superior treatment would cluster near the top achievement level and would exhibit limited interclass variability.

For each Basal versus nonbasal comparison, the lowest class mean and highest class mean on each subtest are reported for each treatment within each project. These class means, based on combined data from the two sexes are recorded in Table 74. There is little to indicate that interclass variability is different for i.t.a. and Basal classrooms. In three of the five projects, the i.t.a. classes exhibited a greater range between the lowest class mean and the highest class mean on the word recognition variable. However, in the other projects greater variability was exhibited by the Basal classrooms. A somewhat similar situation existed with respect to the Paragraph Meaning subtest. The Fry and Hahn projects found greater variability among i.t.a. classrooms, chiefly because of very low achievement on the part of one classroom in the i.t.a. treatment. In each of these projects, the lowest mean achievement was produced by an i.t.a. class. The Hayes project found practically identical interclass variability, although in this project both the lowest and highest i.t.a. class means were considerably above their Basal counterparts. In the last two projects, slightly higher variability was exhibited by the Basal classrooms. A similar lack of consistency is found for the Spelling and Word Study Skills subtests. In certain projects, greater interclass variability was found for the Basal treatment while, in other projects, greater interclass variability was obtained for the i.t.a. classes. Overall, there is no evidence that either the Basal or the i.t.a. treatment tends to increase or decrease interclass variability. Of course, the information in Table 74 says nothing about intraclass variability.

The Basal versus Basal plus Phonics comparison leads to a similar conclusion. There is no clear-cut evidence that interclass variability is greater or smaller in either of the two treatments. The same could be said of the Basal versus Linguistic and Basal versus Phonic/Linguistic comparisons.

The Basal versus Language Experience comparison presents a somewhat different picture. Although not all of the differences go in the same direction, there is an indication that the range between the highest and lowest average class achievement is greater in the Language Experience approach than the Basal program. Note, for example, the tremendous interclass variability for the

Table 74 Ranges of class means of four Stanford subtests by treatment and project

		Word Reading			Paragraph Meaning			Spelling			Word Study Skills		
Project	Treatment	Min.	Max.	Range	Min.	Max.	Range	Min.	Max.	Range	Min.	Max.	Range
Basal vs. i.t.a.													
Fry	Basal	16	22	6	17	23	6	8	13	5	28	37	9
	i.t.a.	14	23	9	10	22	12	4	10	6	30	40	10
Hahn	Basal	18	27	9	17	29	12	8	15	7	29	45	16
	i.t.a.	18	32	14	11	32	21	5	17	12	33	47	14
Hayes	Basal	12	22	10	9	26	17	2	13	11	25	43	18
	i.t.a.	19	32	13	16	34	18	11	19	8	35	49	14
Mazurk	Basal	12	27	15	9	30	21	5	17	12	25	45	20
	i.t.a.	14	27	13	10	28	18	3	13	10	26	43	17
Tanyzer	Basal	13	21	8	12	20	8	6	13	7	28	37	9
	i.t.a.	22	28	6	20	27	7	10	14	4	39	49	10
Basal vs. Basal plus Phonics													
Bordeaux	Basal	16	23	7	14	28	14	5	16	11	27	40	13
	B + P	16	23	7	14	24	10	4	14	10	27	35	8
Hayes	Basal	12	22	10	9	26	17	2	13	11	25	43	18
	B + P	15	28	13	14	29	15	5	17	12	30	43	13
Manning	Basal	11	20	9	8	20	12	1	7	6	19	33	14
	B + P	18	24	6	13	23	10	6	12	6	29	37	8
Murphy	Basal	13	28	15	10	26	16	2	16	14	27	43	16
	B + P	16	31	15	12	29	17	4	17	13	27	49	22
Basal vs. Language Experience													
Cleland	Basal	19	27	8	19	31	12	10	19	9	35	44	9
	LE	14	30	16	19	32	13	6	19	13	31	48	17
Hahn	Basal	18	27	9	17	29	12	8	15	7	29	45	16
	LE	19	28	9	16	31	15	10	18	8	32	45	13
Kendrick	Basal	12	25	13	11	28	17	3	17	14	19	45	26
	LE	15	24	9	8	25	16	4	15	11	25	42	17
Stauffer	Basal	12	19	7	10	20	10	5	13	8	26	41	15
	LE	8	30	22	5	31	26	0	18	18	20	51	31
Basal vs. Linguistic													
Ruddell	Basal	12	22	10	8	24	16	4	12	8	26	37	11
	Ling	14	24	10	11	22	11	3	13	10	25	39	14
Schneyer	Basal	8	29	21	4	32	28	1	18	17	18	49	31
	Ling	6	27	21	5	30	25	0	16	16	17	46	29
Sheldon	Basal	17	25	8	16	27	11	6	17	11	33	46	13
	Ling	12	29	17	8	29	21	3	17	14	28	51	23
Basal vs. Phonic/Linguistic													
Hayes	Basal	12	22	10	9	26	17	2	13	11	25	43	18
	P/L	17	33	16	12	34	22	6	17	11	29	47	18
Tanyzer	Basal	13	21	8	12	20	8	6	13	7	28	37	9
	P/L	24	31	7	19	38	9	12	16	4	37	48	11
Wyatt	Basal	17	26	9	12	30	18	7	18	11	28	50	22
	P/L	22	32	10	18	31	13	11	18	7	36	50	14

Language Experience approach in the first and fourth projects listed in Table 74. In these two projects, the range for each subtest was greater in the Language Experience approach and, usually, the difference between the ranges for the Language Experience and the Basal approach is quite striking. Furthermore, in these two projects, for each subtest the lowest class mean was found in the Language Experience treatment as was the highest class mean. Perhaps this indicates that certain teachers find it difficult to put into practice an instructional program which does not use a structured set of materials which systematically introduces to the child the basic reading skills. On the other hand, this finding might also indicate that certain other teachers find that they can proceed much more efficiently without the "lock-step" inherent, to some extent, at least in the teacher's use of most basal reader programs.

One obvious finding of this part of the analysis is that large differences exist in the mean achievement of various classrooms, even within a treatment and within a project. It is often said that greater differences exist among classrooms within a treatment than between treatments. There is much to support that statement here.

The second approach to assessing variability within treatments was to examine the standard deviations obtained for each treatment on each achievement measure. A relatively large standard deviation for any treatment might indicate that that treatment encouraged superior pupils to achieve up to their capabilities or that the treatment was relatively ineffective for pupils experiencing difficulty in beginning reading. Certainly the variability of pupils taught by a specific method or program is a matter of interest. Table 75 records the standard deviation of scores by treatment on all outcome measures. The standard deviations reported are pooled estimates based on all pupils within classes labeled Basal, i.t.a., Linguistic, Language Experience, Basal plus Phonics, and Phonic/Linguistic. These measures of variability should

Table 75 Standard Deviations[a] of Stanford measures according to treatment

	Word Reading	Paragraph Meaning	Vocabulary	Spelling	Word Study Skills
Basal (4,266)[b]	6.03	7.61	5.72	5.36	8.71
Basal plus Phonics (1,104)	6.49	8.34	5.58	5.34	9.32
i.t.a. (1,055)	6.98	9.54	5.89	5.13	8.96
Language Experience (1,431)	6.60	8.53	5.86	4.85	8.49
Linguistic (760)	6.29	7.33	5.15	4.79	7.76
Lippincott (488)	6.44	9.33	6.08	5.49	8.98

[a] These standard deviations are pooled estimates based on all pupils within class labeled Basal, i.t.a., etc

[b] Indicate number of individuals on whom standard deviation is calculated

Table 76 Comparison of Basal and Nonbasal subjects on the Stanford Word Reading, Paragraph Meaning, and Spelling subtests

		Word Reading			Paragraph Meaning			Spelling		
Methods compared	Number of studies	Nonbasal method superior*	No significant difference	Basal Reader superior*	Nonbasal method superior*	No significant difference	Basal Reader superior*	Nonbasal method superior*	No significant difference	Basal Reader superior*
i.t.a. vs. Basal	5	3	2	0	1	4	0	1	1	3
Linguistic vs. Basal	3	1	2	0	0	3	0	1	2	0
Basal + Phonics vs. Basal	4	0	4	0	1	3	0	1	3	0
Language Experience vs. Basal	4	2	2	0	1	2	1	0	4	0
Phonic/Linguistic vs. Basal	3	3	0	0	2	1	0	1	2	0

*P < .05

Table 77 Comparison of Basal and Nonbasal subjects on the Stanford Vocabulary and Word Study Skills subtests

		Vocabulary			Word Study Skills		
Methods compared	Number of studies	Nonbasal method superior*	No significant difference	Basal Reader superior*	Nonbasal method superior*	No significant difference	Basal Reader superior*
i.t.a. vs. Basal	5	0	5	0	2	3	0
Linguistic vs. Basal	3	0	3	0	1	2	0
Basal + Phonics vs. Basal	4	1	3	0	2	2	0
Language Experience vs. Basal	4	1	3	0	1	3	0
Phonic/Linguistic vs. Basal	3	0	3	0	1	2	0

*P < .05

Table 78 Comparison of Basal and Nonbasal subjects on the Fry and Gates Word Lists

		Fry			Gates		
Methods compared	Number of studies	Nonbasal method superior*	No significant difference	Basal Reader superior*	Nonbasal method superior*	No significant difference	Basal Reader superior*
i.t.a. vs. Basal	5	4	1	0	3	2	0
Linguistic vs. Basal	3	2	1	0	0	3	0
Basal + Phonics vs. Basal	4	1	3	0	2	2	0
Language Experience vs. Basal	3	1	2	0	1	2	0
Phonic/Linguistic vs. Basal	3	3	0	0	3	0	0

*P < .05

be considered illustrative only because of the problems involved in pooling data from different projects. However, there is an unusual degree of similarity in variability among the various treatments. All of the standard deviations for the Word Reading subtest, for example, are greater than six but less than seven. These are certainly negligible differences. The variability is somewhat greater for the Paragraph Meaning subtest. For this test, the lowest variability was found for the Linguistic treatment, while the highest variability was reported for the i.t.a. treatment. However, the differences are probably of limited practical significance. The six standard deviations are within one point of each other for the Vocabulary subtest. The same can be said for the Spelling subtest. Furthermore, the intertreatment variability is only slightly greater for the Word Study Skills subtest. Based on the information reported here, there is little to indicate that the variability of pupils differs to any extent from treatment to treatment.

Summary of treatment comparisons

A summary of the within-projects treatment comparisons is presented in Tables 76, 77, 78, and 79. Each of the tables lists the methods compared and the number of projects in which significant differences favored either the basal approach or the nonbasal approach. The number of projects in which no differences between treatments were found is also recorded. Tables 76 and 77 report significant differences for the Stanford Achievement Tests; Tables 78 and 79 report data on the individual sample measures. As a general finding, it can be stated that the nonbasal programs tended to produce pupils with better word recognition skills than did the basal programs. This finding was especially true with respect to the i.t.a., Phonic/Linguistic, and Basal plus Phonics programs. Differences between basal and nonbasal approaches were less consistent with respect to Paragraph Meaning, Spelling, rate of reading, and reading accuracy. Furthermore, there was little evidence that any approach increased or decreased variability of pupil achievement in reading.

Another general finding is that girls tended to have a greater degree of readiness for reading at the beginning of the first grade and tended to achieve at a higher level of reading at the end of the first grade. In most cases, differences in reading achievement which favored girls at the end of the year disappeared when the crite-

Table 79 Comparison of Basal and Nonbasal subjects on the Gilmore Accuracy and Rate Scores

		Gilmore Accuracy			Gilmore Rate		
Methods compared	Number of studies	Nonbasal method superior*	No significant difference	Basal Reader superior*	Nonbasal method superior*	No significant difference	Basal Reader superior*
i.t.a. vs. Basal	5	1	4	0	0	5	0
Linguistic vs. Basal	3	0	3	0	0	1	2
Basal + Phonics vs. Basal	4	0	3	1	0	4	0
Language Experience vs. Basal	3	1	2	0	0	3	0
Phonic/Linguistic vs. Basal	3	1	2	0	0	3	0

*$P < .05$

rion scores were adjusted for differences in prereading ability. This finding supports the general conclusion that girls are more mature in the first grade and more able to profit from instruction. A related finding in this investigation was that none of the treatments had a unique effect on the achievement of boys and girls. That is, no significant sex by treatment interactions were found to exist. On the average, girls tended to achieve at a higher rate in all programs.

One of the most striking findings was the persistence of project differences in reading achievement even after adjustments were made for differences in pupil readiness for reading. Evidently, reading achievement is influenced by factors peculiar to school systems over and above differences in measured prereading capability of pupils.

Limitations of the analysis of treatment comparisons

There are a number of limitations involved in interpreting the findings of the analysis of treatment comparisons. A first limitation is that not all treatments were represented in all projects. This made it impossible to make direct comparisons between such treatments as i.t.a., Language Experience, Linguistic, and Phonic/Linguistic. The tremendous project differences in achievement would have made comparisons between treatments found in different projects meaningless. As a result, it was only possible to compare the various experimental treatments with the Basal treatment in each project. Of course, these comparisons have been made in the reports of the individual projects.

Another limitation is that treatments labeled Linguistic, Basal, Basal plus Phonics, and i.t.a. did not follow exactly the same program in each project. For example, the Basal reader approach was considered a single treatment even though a variety of Basal programs were used in the various projects. Furthermore, the Linguistic, Basal plus Phonics, and i.t.a. treatments also used different materials from project to project. Furthermore, the Language Experience approach was not exactly the same instructional program in the four projects which utilized this treatment. Grouping programs and materials into a single category should not disguise the fact that actual differences existed in the instructional program within a category.

Another limitation of the study is that there was evidence of nonrandom assignment of pupils to treatments in certain projects. In some cases, there were substantial differences in pupil performance on premeasures for the experimental treatments. The analysis of covariance was used to adjust for premeasure differences, but there is a question of how adequately this statistical technique adjusts for differences in capabilities between groups.

Another limitation which might influence the results is that there appeared to be differences among projects in the extent to which the Hawthorne effect was controlled. It is likely that the newer programs profited from the increased motivation, the greater teacher and parental interest, the awareness on the part of pupils and teachers that experimentation was going on, and similar factors usually associated with new methodological techniques. The extent to which these extraneous factors were controlled in the various projects undoubtedly influenced the results. In this regard, it is likely that the less traditional instructional programs profited from whatever Hawthorne effect was present in the investigation.

Analysis of treatment by readiness level

That part of the analysis which was designed to test for differential treatment effects for pupils who possessed different prereading characteristics is discussed below. The analysis of general treatment effects across all levels of readiness has been presented earlier. The analysis reported presently, however, assesses the relative effectiveness of treatments for pupils of low, average, and high readiness for reading as measured by tests of intelligence, auditory discrimination, and letter knowledge. Using the Basal versus nonbasal comparisons employed in the main analysis reported earlier, subjects were blocked in turn according to performance on the Pintner-Cunningham Primary Test, the Murphy-Durrell Phonemes Test, and the Murphy-Durrell Letter Names Test. A separate analysis of variance was conducted to test differential treatment effects for various levels of performance on the three variables.

The focus of interest in each analysis of variance is the appropriate treatment by intelligence, treatment by auditory discrimination, and treatment by letter knowledge interaction. A significant interaction would indicate that treatments were not operating in the same fashion across all levels of performance on the specific measure being utilized. When a significant interaction of this sort was located, the data were scanned to find the reason. An interesting possible explanation would be that one treatment was more effective for low-readiness pupils; another treatment, more effective for high-readiness pupils.

Because of the nature of this section of the analysis, individuals were used as the experimental unit. Half class means computed separately for the sexes were used as the experimental unit in the analysis of method discussed in the preceding section. Since the same pupils were involved in both analyses, it was possible to compare treatment effects in the two methods of analysis. This comparison is reported in the final report of the Coordinating Center (Bond & Dykstra, 1967, Chapter 8).

Blocking on intelligence

Subjects were categorized according to performance on the Pintner-Cunningham Primary Test in order to test for differential treatment effects for pupils with varying levels of intelligence. Four levels of intelligence were established. Cutoff points were selected in such a way that approximately one fourth of the population in each Basal versus nonbasal comparison fell in each of the four mental age categories. Furthermore, cutoff points were set so as to be identical for all Basal versus nonbasal comparisons. For example, the high intelligence group in all comparisons (Basal versus i.t.a., Basal versus Basal plus Phonics, Basal versus Language Experience, Basal versus Linguistic, and Basal versus Phonic/Linguistic) was composed of those pupils who obtained raw scores of 44 or higher on the intelligence measure. Likewise, the low intelligence group in each of the comparisons consisted of pupils who scored 33 or lower. Pupils in the high-middle range of intelligence scored 39 to 43, while pupils in the low-middle range scored 34 to 38.

Table 80 Cell frequencies for each level of intelligence for the Basal vs. i.t.a. treatments

Project	Sex	Treatment	L (33 or less)	LM (34-38)	HM (39-43)	H (44 or more)
Fry	Male	Basal	17	7	10	21
		i.t.a.	21	13	22	15
	Female	Basal	4	5	16	30
		i.t.a.	12	11	20	20
Hahn	Male	Basal	28	26	44	50
		i.t.a.	29	23	36	44
	Female	Basal	24	19	45	40
		i.t.a.	25	20	31	47
Hayes	Male	Basal	25	8	7	5
		i.t.a.	24	8	2	10
	Female	Basal	16	8	6	12
		i.t.a.	26	10	10	6
Mazurkiewicz	Male	Basal	31	30	52	56
		i.t.a.	44	38	31	67
	Female	Basal	18	22	45	63
		i.t.a.	30	24	38	58
Tanyzer	Male	Basal	38	41	20	19
		i.t.a.	26	32	32	34
	Female	Basal	23	23	27	37
		i.t.a.	12	14	45	47

Table 81 Cell frequencies for each level of intelligence for the Basal vs. Basal plus Phonics treatment

Project	Sex	Treatment	L (33 or less)	LM (34-38)	HM (39-43)	H (44 or more)
Bordeaux	Male	Basal	6	11	14	32
		B+P	5	14	15	31
	Female	Basal	4	3	9	32
		B+P	5	13	15	21
Hayes	Male	Basal	25	8	7	5
		B+P	32	7	5	6
	Female	Basal	16	8	6	12
		B+P	25	13	12	3
Manning	Male	Basal	75	37	31	15
		B+P	64	44	14	12
	Female	Basal	60	41	34	17
		B+P	50	32	27	17
Murphy	Male	Basal	53	17	23	17
		B+P	129	57	60	37
	Female	Basal	42	25	24	13
		B+P	80	46	62	47

Table 82 Cell frequencies for each level of intelligence for the Basal vs. Language Experience treatment

Project	Sex	Treatment	L (33 or less)	LM (34-38)	HM (39-43)	H (44 or more)
Cleland	Male	Basal	68	46	44	34
		LE	56	25	26	24
	Female	Basal	44	42	44	54
		LE	47	32	37	40
Hahn	Male	Basal	28	26	44	50
		LE	21	34	38	42
	Female	Basal	24	19	45	40
		LE	24	20	32	58
Kendrick	Male	Basal	119	65	75	95
		LE	75	61	88	108
	Female	Basal	65	67	75	91
		LE	52	60	82	111
Stauffer	Male	Basal	68	18	25	16
		LE	47	12	26	36
	Female	Basal	21	22	27	22
		LE	41	17	23	36

The numbers of pupils who fell in each of the four levels of intelligence within each sex, treatment, and project are reported in Tables 80, 81, 82, 83, and 84. It should be noted that relatively small numbers of subjects were found in certain cells, a case in point being high intelligence males in the Phonic/Linguistic treatment in Hayes' project. As a result, the findings of this section of the analysis should be interpreted with caution.

Cell frequencies for the Language Experience versus Basal comparison are illustrative of those for other treatment comparisons. These cell frequencies are reported in Table 82. Adding across projects reveals that there were

Table 83 Cell frequencies for each level of intelligence for the Basal vs. Linguistic treatment

Project	Sex	Treatment	L (33 or less)	LM (34-38)	HM (39-43)	H (44 or more)
Ruddell	Male	Basal	18	10	15	17
		Linguistic	25	6	5	15
	Female	Basal	17	14	9	20
		Linguistic	27	7	5	7
Schneyer	Male	Basal	100	16	22	32
		Linguistic	108	22	26	23
	Female	Basal	73	26	25	40
		Linguistic	102	18	22	26
Sheldon	Male	Basal	8	20	19	26
		Linguistic	47	24	40	43
	Female	Basal	7	6	19	38
		Linguistic	43	13	38	68

Table 84 Cell frequencies for each level of intelligence for the Basal vs. Phonic/Linguistic treatment

Project	Sex	Treatment	L (33 or less)	LM (34-38)	HM (39-43)	H (44 or more)
Hayes	Male	Basal	25	8	7	5
		P/L	33	10	6	4
	Female	Basal	16	8	6	12
		P/L	22	9	7	8
Tanyzer	Male	Basal	38	41	20	19
		P/L	22	21	24	39
	Female	Basal	23	23	27	37
		P/L	10	8	23	39
Wyatt	Male	Basal	15	9	32	31
		P/L	14	13	28	34
	Female	Basal	12	25	35	51
		P/L	11	13	44	46

1,431 pupils in the Language Experience group and 1,523 pupils in the Basal group. Adding across sex reveals that there were 1,540 boys and 1,414 girls in all of the projects. A breakdown of intelligence levels by sex results in the finding that there were 405 boys and 452 girls in the highest level of intelligence. The high-middle range of intelligence included 366 boys and 365 girls; the low-middle range of intelligence, 287 boys and 279 girls; and the lowest level of intelligence, 482 boys and 318 girls.

It is possible to analyze Table 82 further to determine the number of subjects who made up the high, high-middle, low-middle, and low intelligence groups. Summing across sex, treatment, and project reveals that there were 857 pupils in the high intelligence group, 731 pupils in the high-middle intelligence group, 566 pupils in the low-middle intelligence group, and 800 pupils in the lowest intelligence group. Therefore, for this particular treatment comparison, the cutoff points selected did not succeed in placing one fourth of the pupils within each of the intelligence levels. However, it should be remembered that the cutoff points were selected to divide the total population (all five treatment comparisons combined) into approximately four levels. Therefore, for each of the five treatment comparisons some deviations from this standard resulted.

After the cutoff points were established, a four-way analysis of variance was employed in which pupils were blocked on intelligence, project, treatment, and sex. For each Basal versus nonbasal comparison, an across-project analysis similar to that described in the preceding section was conducted on the assumption that, within each project, treatments were assigned at random to a set of classes. The across-projects analysis would be meaningful only if no treatment by project interactions were found. The existence of significant treatment by project interactions would indicate that treatments were not operating in the same fashion across all projects and that interpretation of any effects involving treatment would then be difficult. A summary of all of the treatment by project interactions for the various Basal versus nonbasal comparisons is recorded in Table 85. It is obvious that treatments did operate differently in various projects and that, therefore, a within-projects analysis was necessary.

A within-projects analysis for each Basal versus Nonbasal treatment comparison was conducted along the lines of the analysis described in the preceding section. The first step in the within-projects analysis was to conduct an analysis of variance blocking on sex, treatment, and intelligence. Primary attention was focused on the treatment and treatment by level of intelligence effects. In the discussion which follows, only these two treatment effects are reported for each project within each treatment comparison.

Following the analysis of variance, an analysis of covariance was also conducted using the readiness premeasures (except for the intelligence test) as covariates. The treatment effects which resulted from this covariance analysis are reported for each treatment comparison. However, the treatment by readiness interaction is not reported. The use of the seven readiness scores as covariates tended to eliminate treatment differences among four levels of intelligence and also practically eliminated intelligence differences. This destroyed the reason for the analysis which was to see whether or not treatments had a differential effect on high and low readiness pupils. Therefore, interpretation of differential treatment effects is based on the within-projects analysis of variance.

Basal versus i.t.a. treatment comparison. A summary of the within-projects analysis of the i.t.a. versus Basal comparison blocking on sex, intelligence, and treatment is reported in Table 86. Only the treatment and treatment by intelligence effects are reported for each project. Columns A of the table report the analysis of variance. Table 86 reports many significant treatment effects, most of which favor the i.t.a. treatment. However, only one significant treatment by intelligence interaction was found. Obviously, the i.t.a. treatment produced somewhat better readers across all levels of intelligence. There is no indication that the Basal and i.t.a. treatments operated differently for pupils with varying degrees of intellectual ability.

Basal versus Basal plus Phonics treatment comparison. The within-projects analysis of the Basal versus Basal plus Phonics comparison is reported in Table 87. The

Table 85 Treatment by project interactions on Stanford measures for Basal vs. Nonbasal treatment comparisons (blocking on intelligence)

	Word Reading		Paragraph Meaning		Vocabulary		Spelling		Word Study Skills	
Comparisons	A[a]	C[b]	A[a]	C[b]	A[a]	C[b]	A[a]	C[b]	A[a]	C[b]
i.t.a. vs. Basal	23.49**	25.68**	13.94**	17.04**	2.78*	.52	45.96**	43.19**	23.31**	21.70**
Degrees of freedom	(4, 1993)	(4, 1986)								
Basal plus Phonics vs. Basal	5.61**	2.76*	3.13*	1.70	9.37**	6.81**	14.36**	10.41**	4.41**	2.41
Degrees of freedom	(3, 1658)	(3, 1651)								
Language Experience vs. Basal	12.45**	10.62**	30.47**	28.54**	6.15**	8.58**	4.08**	3.31*	8.85**	7.61**
Degrees of freedom	(3, 2890)	(3, 2883)								
Linguistic vs. Basal	12.57**	17.32**	6.29**	6.13**	2.80	5.30**	12.07**	12.73**	5.52**	5.67**
Degrees of freedom	(2, 1309)	(2, 1302)								
Phonic/Linguistic vs. Basal	13.11**	12.31**	8.64**	8.80**	6.51**	2.38	5.79**	8.58**	10.33**	9.26**
Degrees of freedom	(2, 965)	(2, 958)								

[a] Reports analysis of variance

[b] Reports analysis of covariance with all seven premeasures as covariates

* Interaction is significant at .05 level

** Interaction is significant at .01 level

BEST COPY AVAILABLE

Table 86 Selected treatment effects from within-projects analysis of variance and covariance on Stanford measures for Basal vs. i.t.a. comparison (blocking on intelligence)

		Word Reading		Paragraph Meaning		Vocabulary		Spelling		Word Study Skills	
P[a]	Effect	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]
1	Treatment	5.09*	12.29**	.54	.00	.12	.61	17.92†	15.61†	5.38*	11.15**
	Treatment × Intelligence	2.33		1.27		.82		.21		.23	
2	Treatment	10.24**	5.48*	.07	.93	1.04	.01	17.06†	32.68†	2.49	.19
	Treatment × Intelligence	2.08		1.08		2.54		1.04		1.08	
3	Treatment	60.84**	60.18**	12.75**	6.76**	1.13	.12	71.39**	69.62**	36.07**	31.07**
	Treatment × Intelligence	.86		2.96††		.12		1.11		2.08	
4	Treatment	3.08	28.31**	.08	11.42**	4.69***	.01	116.92†	81.05†	7.85†	.19
	Treatment × Intelligence	1.04		2.38		.66		.77		.24	
5	Treatment	125.01**	176.45**	55.64**	81.20**	3.51	2.00	3.95*	9.67**	71.39**	90.40**
	Treatment × Intelligence	1.04		1.86		1.92		.17		.57	

[a] Projects in numerical order are Fry, Hahn, Hayes, Mazurkiewicz, and Tanyzer

[b] Reports analysis of variance. Treatment effects are based on 1 and 1993 *df*; interactions, on 3 and 1993 *df*

[c] Reports covariance with all seven premeasures as covariates. Treatment effects based on 1 and 1986 *df*

* Significant difference favoring i.t.a. at .05 level

** Significant difference favoring i.t.a. at .01 level

*** Significant difference favoring Basal at .05 level

† Significant difference favoring Basal at .01 level

†† Interaction significant at .05 level

Table 87 Selected treatment effects from within-projects analysis of variance and covariance on Stanford measures for Basal vs. Basal plus Phonics comparison (blocking on intelligence)

		Word Reading		Paragraph Meaning		Vocabulary		Spelling		Word Study Skills	
P[a]	Effect	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]
1	Treatment	3.40	1.84	.37	.01	3.17	.69	3.21	2.74	1.01	.10
	Treatment × Intelligence	1.54		1.29		.52		.80		.71	
2	Treatment	3.67	9.40**	1.18	3.82	6.62†	5.12†	10.93**	16.29**	3.27	6.42*
	Treatment × Intelligence	1.05		.16		.46		1.29		.19	
3	Treatment	51.48**	17.73**	29.94**	6.76**	41.56**	17.86**	94.42**	52.18**	49.42**	22.07**
	Treatment × Intelligence	.43		.50		.33		.18		.41	
4	Treatment	4.37*	1.62	11.27**	7.63**	15.36**	22.75**	.91	.01	7.76**	5.38*
	Treatment × Intelligence	.80		1.85		.88		.14		.06	

[a] Projects in numerical order are Bordeaux, Hayes, Manning, and Murphy

[b] Reports analysis of variance. Treatment effects are based on 1 and 1658 *df*; interactions, on 3 and 1658 *df*

[c] Reports covariance with all seven premeasures as covariates. Treatment effects are based on 1 and 1651 *df*

* Significant difference favoring Basal plus Phonics at .05 level

** Significant difference favoring Basal plus Phonics at .01 level

† Significant difference favoring Basal at .05 level

BEST COPY AVAILABLE

Table 88 Selected treatment effects from within-projects analysis of variance and covariance on Stanford measures for Basal vs. Language Experience comparison (blocking on intelligence)

		Word Reading		Paragraph Meaning		Vocabulary		Spelling		Word Study Skills	
P[a]	Effect	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]
1	Treatment	17.36**	43.25**	9.39**	22.54**	19.46**	42.52**	1.08	.03	8.28**	23.09**
	Treatment × Intelligence	.43		1.24		3.49††		.46		.81	
2	Treatment	5.97*	5.60	.42	.10	1.43	.96	1.18	.75	.89	.49
	Treatment × Intelligence	1.15		1.95		.68		.46		1.06	
3	Treatment	6.23***	.04	81.88†	52.38†	1.52	.68	5.24***	1.12	17.22†	1.84
	Treatment × Intelligence	2.02		.92		3.92†††		.47		1.91	
4	Treatment	17.86**	20.68**	8.88***	9.65***	.59	.64	5.72*	9.07**	.52	.33
	Treatment × Intelligence	10.40†††		12.58†††		8.22†††		6.34†††		10.82†††	

[a] Projects in numerical order are Cleland, Hahn, Kendrick, and Stauffer

[b] Reports analysis of variance. Treatment effects based on 1 and 2890 *df*; interactions, on 3 and 2890 *df*

[c] Reports covariance with all seven premeasures as covariates. Treatment effects based on 1 and 2883 *df*

* Significant difference favoring Language Experience at .05 level

** Significant difference favoring Language Experience at .01 level

*** Significant difference favoring Basal at .05 level

† Significant difference favoring Basal at .01 level

†† Interaction significant at .05 level

††† Interaction significant at .01 level

treatment effects reported for both the analysis of variance and covariance favor the Basal plus Phonics treatment in every instance except for the Vocabulary subtest in project 2 (Hayes). However, no treatment by intelligence interactions were found to be significant in any of the projects. These two findings would indicate that the Basal plus Phonics approach was somewhat superior to the Basal approach for high intelligence, average intelligence, and low intelligence pupils alike and that the extent of this superiority was consistent across intelligence levels. Furthermore, there was no indication that one treatment was better for high intelligence pupils, or the other treatment better for low intelligence pupils.

Basal versus Language Experience treatment comparisons. A summary of the within-projects analysis for the Basal versus Language Experience treatment comparison is presented in Table 88. Again, only the treatment and treatment by interaction effects are reported. A number of treatment effects were found to be statistically significant in the various projects. In three of the four projects, the differences favored the Language Experience approach, while in project 3 (Kendrick) the differences favored the Basal approach. However, a number of significant treatment by intelligence interactions were found. Projects 1 (Cleland) and 3 (Kendrick) reported significant interactions between treatment and intelligence on the Vocabulary subtest, while project 4 (Stauffer) recorded significant interactions on all five subtests.

The reason for the interaction in project 4 (Stauffer) is reported in Table 89. On each of the subtests, the Language Experience approach was superior to the Basal approach for the upper three levels of intelligence. However, in each case the Basal approach produced higher performance for the low intelligence pupils. The data from project 4 (Stauffer) would indicate that the less capable pupil would profit more from a Basal program, while more capable pupils would profit from the Language Experience approach. However, this conclusion is tempered by the fact that significant project by treatment interactions were not found on the four reading-related achievement measures for the other three projects in the Basal versus Language Experience comparison.

Because the interactions were primarily restricted to one project, the data were examined further. It was possible that an examination of performance on premeasures would reveal the reason for the significant interactions. Table 90 presents relevant information. The same pattern of performance existed on the premeasures as had been found on the Stanford Achievement Test. Within the lowest range of intelligence, Basal pupils scored better on readiness measures than did children in the Language Experience treatment. Within the higher range of intelligence, however, Language Experience pupils were more ready for reading. This suggests that the significant treatment by intelligence interactions on the postinstructional achievement measures was simply a reflection of treatment by project interactions on readiness measures. Low intelligence Basal pupils were more ready for reading in other respects than were low intelligence Language Experience pupils and they became more successful readers. High intelligence Language Experience pupils were more ready for reading in respects other than intelligence than were their high intelligence Basal counterparts and they became better readers. Therefore, the interactions between treatment and intelligence on the achievement measures probably are of little educational significance.

Basal versus Linguistic treatment comparison. The summary of the within-projects analysis for the Basal versus Linguistic treatment comparison is reported in Table 91. The anaylsis of variance and analysis of covariance point out a number of significant treatment differences, some favoring the Basal approach and some favoring the Linguistic approach. However, only one significant treatment by intelligence interaction was found, that for the Vocabulary subtest. Therefore, there is nothing to indicate that the treatments operated differentially for pupils of high or low intelligence. In some projects, the Basal approach was superior and the superiority was evident for all ranges of intelligence. In other projects, the Linguistic approach was superior and the superiority held up across all ranges of intelligence.

Basal versus Phonic/Linguistic treatment comparison. Selected treatment effects from the within-projects

Table 89 Unadjusted Standford means for project four in Basal vs. Language Experience comparison by treatment and level of intelligence

	Word Reading		Paragraph Meaning		Spelling		Word Study Skills	
Level of Intelligence	Basal	LE	Basal	LE	Basal	LE	Basal	LE
High intelligence 44 or more	19.4	25.9	20.3	28.6	12.8	16.3	38.7	45.7
High–middle intelligence 39–43	17.7	22.5	17.9	23.4	10.1	13.7	37.3	39.8
Low–middle intelligence 34–38	16.5	20.7	15.4	19.9	9.8	12.8	34.2	37.4
Low intelligence 33 or less	13.1	11.5	12.7	9.2	6.0	4.3	28.4	23.4

Table 90 Unadjusted premeasure means for project four in Basal vs. Language Experience comparison by treatment and level of intelligence

	Murphy-Durrell Phonemes		Murphy-Durrell Letter Names		Murphy-Durrell Learning Rate	
Level of Intelligence	Basal	LE	Basal	LE	Basal	LE
High intelligence 44 or more	19.2	23.7	20.0	26.1	22.6	24.5
High–middle intelligence 39–43	19.4	22.9	19.9	23.8	22.1	22.8
Low–middle intelligence 34–38	19.2	22.7	18.4	22.1	19.4	21.7
Low intelligence 33 or less	18.0	17.8	18.2	16.3	18.9	17.4

analysis of the Basal versus Phonic/Linguistic comparisons are presented in Table 92. The analysis of variance and analysis of covariance found many significant treatment differences favoring the Phonic/Linguistic treatment. However, only two treatment by project interactions were found to be significant. One of these interactions involved the Vocabulary subtest, while the other involved the Paragraph Meaning subtest. It is apparent that the Phonic/Linguistic program was superior across all levels of intelligence. There was no indication that the Basal approach was better for pupils of a given intellectual capability, while the Phonic/Linguistic program was better for pupils of a different level of intelligence.

Summary. For four of the five Basal versus nonbasal treatment comparisons, there was no evidence of differential treatment effects according to pupil intelligence. Either

Table 91 Selected treatment effects from within-projects analysis of variance and covariance on Stanford measures for Basal vs. Linguistic comparison (blocking on intelligence)

		Word Reading		Paragraph Meaning		Vocabulary		Spelling		Word Study Skills	
P[a]	Effect	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]
1	Treatment	24.92**	44.46**	.04	.33	.33	2.95	7.12**	9.56**	3.06	5.78*
	Treatment × Intelliegence	1.61		.97		.83		1.58		.31	
2	Treatment	1.98	.57	4.43***	2.84	12.34†	10.56†	15.98†	15.02†	9.81†	7.19†
	Treatment × Intelligence	1.06		.41		4.32††		.13		1.00	
3	Treatment	.01	6.77**	28.99†	22.45†	9.89†	1.82	17.69†	11.40†	6.66†	1.06
	Treatment × Intelligence	.99		.76		1.19		.67		.55	

[a] Projects in numerical order are Ruddell, Schneyer, and Sheldon

[b] Reports analysis of variance. Treatment effects based on 1 and 1309 *df*, interactions, on 3 and 1309 *df*

[c] Reports covariance with all seven premeasures as covariates. Treatment effected based on 1 and 1302 *df*

* Significant difference favoring Linguistic at .05 level

** Significant difference favoring Linguistic at .01 level

*** Significant difference favoring Basal at .05 level

† Significant difference favoring Basal at .01 level

†† Interaction significant at .01 level

Table 92 Selected treatment effects from within-projects analysis of variance and covariance on Stanford measures for Basal vs. Phonic/Linguistic comparison (blocking on intelligence)

		Word Reading		Paragraph Meaning		Vocabulary		Spelling		Word Study Skills	
P[a]	Effect	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]	A[b]	C[c]
1	Treatment	64.75**	100.71**	15.85**	27.09**	.08	1.96	25.84**	42.51**	19.94**	39.40**
	Treatment × Intelliegence	1.90		1.11		3.55†		1.90		1.67	
2	Treatment	173.11**	142.23**	44.13**	23.98**	30.49**	7.79**	35.90**	17.76**	64.37**	33.15**
	Treatment × Intelligence	1.45		4.04††		.18		.59		1.28	
3	Treatment	51.01**	65.41**	3.80	3.30	1.12	.28	8.65**	10.77**	6.48*	8.52*
	Treatment × Intelligence	.42		.11		.19		.16		.12	

[a] Projects in numerical order are Hayes, Tanyzer, and Wyatt

[b] Reports analysis of variance. Treatment effects are based on 1 and 958 *df*, interactions, on 3 and 958 *df*

[c] Reports covariance with all seven premeasures as covariates. Treatment effects based on 1 and 951 *df*

* Significant difference favoring Phonic/Linguistic at .05 level

** Significant difference favoring Phonic/Linguistic at .01 level

† Interaction significnat at .05 level

†† Interaction significant at .01 level

BEST COPY AVAILABLE

no difference between the Basal and nonbasal treatment existed, or the superior treatment operated in the same fashion across all levels of intelligence. On the other hand, there was some evidence in the analysis of the Basal versus Language Experience comparison that the Language Experience approach was superior for average and above-average pupils, while the Basal approach was better for the pupils at the lowest level of intelligence. This finding, however, has limited significance in light of the fact that differences between treatments in performance on premeasures other than intelligence were very similar to the differences found on achievement measures.

Blocking on phonemes

Pupils were classified as having high, average, or low auditory discrimination by setting up cutoff points on the Murphy-Durrell Phonemes Test. The cutoff points were established so that approximately one third of the total population fell into each of three categories. Information concerning the numbers of pupils from each of the projects who fell in each of the auditory discrimination levels is reported in the Appendix to the Coordinating Center's final report (Bond & Dykstra, 1967), as are summaries of the across-projects and within-projects analyses. The results can be summarized by stating that, except for the Basal versus Language Experience comparison, very few treatment by auditory discrimination interactions were found to be significant. Where treatment differences were found to be significant in a Basal versus nonbasal comparison, the usual finding was that the superior treatment was superior over all three levels of auditory discrimination. There was no indication that Basal and Nonbasal treatments operated differentially for the three levels of auditory discrimination. Tables similar to those presented for the various treatment comparisons blocked on intelligence are presented in the Appendix (Bond & Dykstra, 1967) for the various treatment comparisons blocked on the Murphy-Durrell Phonemes Test.

Auditory discrimination by treatment interactions were found to be significant in the Language Experience versus Basal comparisons. Pupils with low auditory discrimination profited more from instruction in a Basal program, while pupils with average and high auditory discrimination gained more from a Language Experience approach. This finding, again, must be interpreted in light of an identical treatment by auditory discrimination interaction on many premeasures. Basal pupils in the low auditory discrimination group were superior to similar Language Experience pupils on the Letter Names, Learning Rate, and Pintner-Cunningham Primary Test. The reverse was true of pupils with high auditory discrimination skills. Among this group, the Language Experience pupils were superior in performance on the premeasures. This interaction involving premeasures probably explains the interaction involving postmeasures.

Blocking on letter knowledge

An analysis similar to the one used blocking on intelligence and auditory discrimination was conducted blocking on letter knowledge. Pupils were placed in approximately equal numbers in four categories according to letter knowledge as measured by the Murphy-Durrell Letter Names Test. The numbers of pupils within each project who were placed in the four levels of letter knowledge are presented in the Appendix to the Coordinating Center's final report (Bond & Dykstra, 1967) as are summaries of the analysis of variance and analysis of covariance pertinent to this aspect of the data analysis. In general, the results showed that few, if any, significant treatment by letter knowledge interactions were found for the Basal versus i.t.a., Basal versus Basal plus Phonics, Basal versus Linguistic, and Basal versus Phonic/Linguistic comparisons. Whenever treatment differences were found to be significant, the superior treatment was superior across all levels of letter knowledge. There was no indication that the Basal treatment was particularly effective for pupils with high or low letter knowledge, while the nonbasal treatment was more effective for pupils with the opposite characteristics.

This finding did not hold true for the Basal versus Language Experience comparison. In that treatment comparison, a number of treatment by letter knowledge interactions (mostly in one project) were found to be significant. An examination of the data revealed that these interactions resulted from the fact that pupils low in letter knowledge profited more from the Basal approach, while pupils at the higher levels of letter knowledge achieved better under the influence of the language arts program. However, the same problem of dissimilarity in readiness characteristics of Basal and Language Experience pupils that existed in the analysis blocking on itelligence and the analysis blocking on auditory discrimination was also evident in this case. The significant treatment by letter knowledge interactions on achievement measures appeared to be merely reflections of treatment by letter knowledge interactions on other premeasures.

Summary

For four of the five Basal versus Nonbasal comparisons, there was no evidence of a differential treatment effect for various levels of intelligence, auditory discrimination, or letter knowledge. Very few, if any, significant treatment by intelligence, treatment by auditory discrimination, or treatment by letter knowledge interaction effects were found to be significant. This finding of no interaction between treatment and readiness characteristics generally held true for the Basal versus i.t.a., Basal versus Basal plus Phonics, Basal versus Linguistic, and Basal versus Phonic/Linguistic treatment comparisons.

A somewhat different conclusion could be drawn from the analysis involving the Basal versus Language Experience comparison. For this treatment comparison, a number of treatment by intelligence, treatment by auditory discrimination, and treatment by letter knowledge interactions were found to be significant. The interactions resulted from the fact that the least mature pupils achieved better in a Basal program than in a Language Experience program. Conversely, more capable students with respect to these skills profited more from a Language Experience approach. It should be emphasized again, however, that this finding probably resulted from similar differential readiness characteristics of Basal and Language Experience pupils and, therefore, has questionable significance.

Summary and conclusions

The present study was designed to obtain information relevant to three basic questions: (1) To what extent are various pupil, teacher, class, school, and community characteristics related to pupil achievement in first-grade reading and spelling? (2) Which of the many approaches to intial reading instruction produces superior reading and spelling achievement at the end of the first grade? (3) Is any program uniquely effective or ineffective for pupils with high or low readiness for reading?

Analysis of relationships

Summary of findings. The findings of the investigation relevant to the first question above can be summarized as follows:

(1) The single best predictor of first-grade reading success among the premeasures used in this investigation was the Murphy-Durrell Letter Names Test. This test correlated between .52 and .60 with both the Stanford Word Reading and Stanford Paragraph Meaning subtests for each of the six treatments used in the investigation.

(2) The Murphy-Durrell Phonemes and the Pintner-Cunningham Primary Test also correlated relatively highly with the criterion measures. Each of these tests correlated .40 or greater with both the Word Reading and Paragraph Meaning subtests for each of the six treatments.

(3) The other readiness tests used in this study correlated positively with the reading measures, but to a smaller extent. Correlations with reading were usually .40 or less for these premeasures.

(4) For the subtests with the best predictive ability (Letter Names, Phonemes, Pintner-Cunningham), there was little evidence of differential prediction of reading success in the programs used in this study. Correlations between these premeasures and reading were very similar for pupils in the Basal, i.t.a., Basal plus Phonics, Language Experience, Linguistic, and Phonic/Linguistic groups.

(5) A correlation coefficient of .86 was found between the Fry Test of Phonetically Regular Words and the Gates Word Pronunciation Test for the Basal treatment. Each of these tests was administered individually to a sample, but they differed in the degree to which words were controlled on the basis of sound-symbol regularity. The Fry test consisted of words with high regularity, while the Gates test consisted of words selected on the basis on frequency of usage with no control of sound-symbol relationship. Furthermore, the Word Reading subtest from the group-administered Stanford Achievement Test correlated .72 with the Fry Word List and .78 with the Gates Word Pronunciation Test for the Basal group. Correlations for the treatments other than Basal were very similar.

(6) For the range of class sizes reported in this study, there was a negligible correlation between class size and reading achievement. Furthermore, pupil absence and child age were negatively related to the various reading measures. However, these correlations were also negligible with the largest of them being –.22.

(7) The total experience of teachers correlated between .24 and .34 with the five Stanford Achievement measures. Teacher experience in the first grade correlated between .20 and .30 with the same measures. A rating of general overall teacher efficiency correlated between .10 and .22 with the five achievement measures.

(8) The accuracy score on the Gilmore Oral Reading Test correlated between .81 and .90 with the Gates Word Pronunciation Test for the various reading programs.

(9) The Stanford Word Reading Test, a measure of word recognition, and the Stanford Paragraph Meaning Test, a measure of comprehension, correlated between .71 and .83 for the various programs.

Conclusions. From the correlation relationships found in the study, the following conclusions can be drawn:

(1) There are many pupil capabilities related to the success children have in beginning reading. The results of this study would indicate that a fair amount of the variation in pupil success can be accounted for by the attributes brought to the learning situation. Such pupil capabilities as auditory and visual discrimination and pre-first-grade familiarity with print and intelligence, are all substantially related to success in learning to read under whatever approach to initial instruction is used.

(2) Among those attributes measured in this study, knowledge of letter names and the ability to discriminate

between word sounds appear to have the greatest relationship to reading success under each of the various methods of instruction employed. The knowledge of letter names gained prior to initial instruction alone would account for approximately 25 to 36 percent of the variation in reading ability found at the end of the year under the various methods of instruction used for this study. It is also interesting to note that the predictive validity of a single subtest such as the Letter Names subtest is of approximately the same magnitude as the predictive validity of an entire reading readiness battery. Therefore, it is probably not necessary to give a complete readiness test if prediction of reading success is the only objective.

(3) Test constructors should note the high positive relationship between the length of the various readiness tests, their variability, and their predictive validity. The length of the tests alone might account for the differential predictive power found among the premeasures used in this study.

(4) Although there were some differences in the magnitude of correlations between prereading pupil characteristics and success in reading under the various methods, there was enough uniformity in these relationships to conclude that no one method of instruction would uniquely overcome the limitations imposed on children by deficiencies in any characteristic measured in this study. Furthermore, no approach used was found to be uniquely effective for pupils who scored well on any of the premeasures.

(5) There were some differences in the magnitude of the correlations between the pretest and the Stanford Paragraph Meaning and Stanford Word Reading tests. However, there was enough uniformity in these relationships to conclude that no one attribute measured by the premeasures would predispose the child to having specific difficulty in word recognition as opposed to comprehension or vice versa. In this respect, it should be noted that the correlation between the two reading tests was so high that little differential prediction could be expected.

(6) Although no relationship between class size and success in first-grade reading was found in this study, the conclusion that class size makes no difference would be unwarranted. No very large or very small classes were represented in the study. It can be concluded, however, that the addition or subtraction of a pupil or two would not seriously influence successful teaching.

(7) The relationship between pupil age and reading success indicates that the younger child did somewhat better than did his older classmate. The correlation was so small that differences in age of these pupils accounted for little, if any, of the differences in reading success. Furthermore, the negative relationship between achievement and chronological age may be a result of selective admission procedures whereby unusually mature children are admitted to school at a young age.

(8) A child who has the ability to read phonetically regular words also has skill in reading words of high utility even though these latter words may be highly irregular. Similarly, children who can read words orally in individual test situations also are relatively successful in reading words silently in a group testing situation. Therefore, in most instances, it is probably not necessary to employ both individual and group measures of word recognition. Similarly, it is probably not necessary to use different tests to evaluate the reading ability of pupils who learn to read by means of linguistic programs (where words are controlled on the basis of sound-symbol regularity) from those used to test reading ability of pupils from basal reading programs.

(9) The high intercorrelations found in the variety of reading measures used in this study indicate that reading at the end of the first grade is largely a unitary accomplishment depending upon the ability to recognize words accurately and to associate meaning with those words. An alternative conclusion, of course, is that it is difficult to develop tests which differentiate abilities at this early level even if they exist.

(10) From the correlation studies, the evidence is that teacher experience and efficiency ratings are only slightly related to pupil success. While there is ample evidence that class differences influence reading success, the estimates of teacher efficiency used in this study did not explain these differences.

Analysis of methodology

The relative effectiveness of the various instructional programs used in this investigation was evaluated in two different ways. The major technique was to compare various Nonbasal programs with Basal programs used in the same project. The newer experimental programs were evaluated by comparing their relative effectiveness with that of the well-known Basal reading programs. This analysis was considered the appropriate one to be used in the study. However, an analysis was also conducted whereby each treatment within each project was compared with all of the other treatments in all of the other projects. In this latter analysis, pupil differences in readiness among the various treatments and projects were adjusted by means of covariance as were teacher differences in experience. Because of tremendous project differences in achievement even after teacher and pupil characteristics had been controlled statistically, this method of analysis was conducted for informational purposes only and is presented in the report of the Coordinating Center (Bond & Dykstra, 1967, Chapter 9).

Summary of findings from Basal versus Nonbasal comparisons. The findings of the Basal versus i.t.a., Basal versus Basal plus Phonics, Basal versus Language Experience, Basal versus Linguistic, and Basal versus Phonic/Linguistic treatment comparisons are summarized below.

(1) *Summary of Basal versus i.t.a. comparisons.* The i.t.a. and Basal approaches were of approximately equal effectiveness in terms of pupils' achievement on the Paragraph Meaning test. However, the i.t.a. treatment produced superior word recognition abilities as measured by the Word Reading subtest of the Stanford and the Fry and Gates word lists. Evidence concerning the spelling ability of pupils in the two groups was inconclusive. The Basal subjects were superior in spelling ability in three projects, but the i.t.a. subjects were superior in a fourth project. No differences were found between treatments in reading accuracy and rate as measured by the Gilmore Oral Reading Test.

(2) *Summary of Basal versus Basal plus Phonics comparison.* In general, Basal programs accompanied by supplementary phonics materials produced significantly greater achievement in reading than did Basal materials alone. This superiority was especially pronounced in the across-projects analysis of mean performance on the Stanford Achievement tests and the Fry and Gates word recognition tests. Practically all differences on these measures favored the Basal plus Phonics group, even though some of the differences failed to reach statistical significance. No differences in rate or accuracy of oral reading were found between the two treatments.

(3) *Summary of Basal versus Language Experience comparison.* Relatively few significant differences were found between the Language Experience and Basal approaches. Those significant differences which were found to exist generally favored the Language Experience approach. However, these sporadic differences were often not of much practical significance in terms of actual reading achievement.

(4) *Summary of Basal versus Linguistic.* The most common finding for the Linguistic versus Basal comparison in the various projects was that of no difference between treatments. The Linguistic group tended to outperform the Basal group on tests of word recognition, while the Basal group exhibited somewhat greater speed and accuracy in reading. No differences in comprehension were ascertained.

(5) *Summary of Basal versus Phonic/Linguistic comparison.* The Phonic/Linguistic program was superior to the Basal program utilized in the projects of this investigation. The Phonic/Linguistic program produced pupils with superior word reading, paragraph meaning, spelling, and word study skills. Phonic/Linguistic pupils were also superior on the Fry Test of Phonetically Regular Words and the Gates Word Pronunciation Test. No significant differences were found in rate or accuracy of oral reading.

In general, there was less difference in variability among treatments than in mean achievement among treatments. Standard deviations on each of the outcome measures were very similar for the Basal, i.t.a., Basal plus Phonics, Language Experience, Linguistic, and Phonic/Linguistic pupils. Furthermore, the interclass variation within the various treatments was very similar except for the Language Experience approach. Wide differences in mean achievement of classrooms were found for all of the programs. However, the range between the highest and lowest average class achievement in the Language Experience approach generally was greater than the range for the Basal classrooms in the same project.

Another general finding was that girls tended to have a greater degree of readiness for reading at the beginning of first grade and tended to read at a higher level of reading at the end of the first grade. In most cases, differences in reading achievement which favored girls at the end of the year disappeared when criterion scores were adjusted for differences in prereading ability. A related finding in this investigation was that none of the treatments had a unique effect on the achievement of boys and girls. That is, no significant sex by treatment interactions were found to exist. On the average, girls tended to be better readers in all programs.

One of the most striking findings was the persistence of project differences in reading achievement, even after adjustments were made statistically for differences in pupil readiness for reading. Evidently, reading achievement is influenced by factors peculiar to school systems over and above differences in prereading capabilities of pupils.

One other common finding was that statistically significant treatment by project interactions were found in most of the Basal versus Nonbasal comparisons. In general, treatments did not operate in the same fashion across projects.

Conclusions. The findings of the analysis of methodology led to the following conclusions:

(1) Word study skills must be emphasized and taught systematically regardless of what approach to initial reading instruction is used.

(2) Combinations of programs, such as a basal program with supplementary phonics materials, often are superior to single approaches. Furthermore, the success of such methods as the Language Experience approach indicates that the addition of language experiences to any kind of reading program can be expected to make a contribution.

(3) Innovative programs such as Linguistic readers are especially effective in the word recognition area. The superiority of these programs to basal programs is not as evident in the area of comprehension. It is likely that basal programs should develop a more intensive word study skills element, while programs which put major emphasis on word recognition should increase attention paid to other reading skills.

(4) It is necessary for teachers to make differential expectations concerning mean achievement of boys and girls. On the average, boys cannot be expected to achieve at the same level as girls, at least with the materials, methods, and teachers involved in this investigation. A probable explanation from the data of this study is that boys are less ready to read when they enter school.

(5) Boys and girls do not profit uniquely from any of the programs utilized in this investigation. On the average, girls' achievement is superior to boys' no matter what approach to beginning reading is used.

(6) Reading programs are not equally effective in all situations. Evidently, factors other than method, within a particular learning situation, influence pupil success in reading.

(7) Reading achievement is related to other characteristics in addition to those investigated in this study. Pupils in certain school systems became better readers than pupils in other school systems even when pupil characteristics were controlled statistically. Furthermore, these differences in achievement from project to project do not seem to be directly related to the class, school, teacher, and community characteristics appraised in this study.

(8) Pupils taught to read by means of a transitional alphabet such as i.t.a. may experience greater difficulty making the transition to traditional orthography in spelling than they do in reading. Longitudinal information is necessary to study this problem.

(9) Future research might well center on teacher and learning situation characteristics rather than method and materials. The tremendous range among classrooms within any method points out the importance of elements in the learning situation over and above the methods employed. To improve reading instruction, it is necessary to train better teachers of reading rather than to expect a panacea in the form of materials.

(10) Children learn to read by a variety of materials and methods. Pupils become successful readers in such vastly different programs as the Language Experience approach with its relative lack of structure and vocabulary control and the various Linguistic programs with their relatively high degree of structure and vocabulary control. Furthermore, pupils experienced difficulty in each of the programs utilized. No one approach is so distinctly better in all situations and respects than the others that it should be considered the one best method and the one to be used exclusively.

(11) The expectation of pupil accomplishment in initial reading instruction probably should be raised. Programs which introduced words at a more rapid pace tended to produce pupils with superior word recognition abilities at the end of the first grade. Children today tend to be better equipped for reading instruction when they enter first grade than they were some years ago and they are probably prepared to learn more words and develop more mature study skills than are currently expected of them in many programs.

(12) Indications are that the initial reading vocabulary should be selected with a greater balance between phonetically regular words and high-utility words. It is likely that introducing words solely on the basis of frequency of use presents an unusually complex decoding task for the beginning reader. On the other hand, it appears that presenting only phonetically regular words makes it very difficult to write meaningful material.

(13) A writing component is likely to be an effective addition to a primary reading program. In the first place, the Language Experience approach, which involves considerable written expression, was an effective program of instruction. In addition, programs such as i.t.a. and Phonic/Linguistic, both of which were relatively effective, encourage pupils to write symbols as they learn to recognize them and to associate them with sounds. This appears helpful to the pupil in learning sound-symbol relationships. Furthermore, it is likely that writing such common, but irregular, words as *the* helps the child commit them to his sight vocabulary.

(14) It is impossible to assess the relative effectiveness of programs unless they are used in the same project. Project differences are so great, even when pupil readiness for reading is controlled, that a program utilized in a favored project would demonstrate a distinct advantage over one used in a less favored project regardless of the effectiveness of the program.

(15) The relative success of the Nonbasal programs compared to the basal programs indicates that reading instruction can be improved. It is likely that improvement would result from adopting certain elements from each of the approaches used in this study. The first step would be to determine the elements within the various approaches most important to the success of that program. For example, the i.t.a. and Phonic/Linguistic programs, both of which were relatively effective, have in common a vocabulary controlled on sound-symbol regularity, introduction of a relatively large reading vocabulary, and emphasis on writing symbols as a means of learning them. It would be interesting to know which of these ele-

ments, if any, are primarily responsible for the effectiveness of the program. Perhaps an instructional program which incorporated the most important elements of all of the approaches used in the study would be a more effective method of teaching than any currently in use.

Analysis of treatment by readiness level

In this section of the analysis, pupils were blocked in turn according to levels of ability as measured by an intelligence test, an auditory discrimination test, and a test of letter knowledge. Interactions between treatments and each of these readiness measures were examined to determine whether or not there was a differential treatment effect for pupils of varying levels of readiness.

Summary of findings. For four of the five Basal versus Nonbasal comparisons, there was no evidence of differential treatment effects for various levels of intelligence, auditory discrimination, or letter knowledge. Very few, if any, treatment by intelligence, treatment by auditory discrimination, or treatment by letter knowledge interaction effects were found to be significant. This finding of no interaction between treatment and readiness characteristics generally held true for the Basal versus i.t.a., Basal versus Basal plus Phonics, Basal versus Linguistic, and Basal versus Phonic/Linguistic comparisons.

A somewhat different situation existed for the Basal versus Language Experience comparison. For this treatment comparison, a number of treatment by intelligence, treatment by auditory discrimination, and treatment by letter knowledge interactions were found to be significant. The interactions resulted from the fact that the least mature pupils achieved better in a Basal program than in a Language Experience approach, while more capable students with respect to these skills profited more from a Language Experience approach. However, this finding was tempered by the fact that the low readiness Basal pupils were generally superior to the low readiness Language Experience pupils on premeasures other than the one used for blocking. It was not surprising to find that they were superior in achievement. On the other hand, the high readiness Basal pupils were inferior to the high readiness Language Experience pupils on premeasures other than those used for blocking. It is possible that the treatment by readiness interaction on the achievement measures was primarily a result of similar interaction on the premeasures.

Conclusions. With respect to the i.t.a. versus Basal, Basal plus Phonics versus Basal, Linguistic versus Basal, and Phonic/Linguistic versus Basal treatment comparisons, the following conclusions can be drawn:

(1) Programs which were superior in the various Basal versus Nonbasal comparisons tended to be superior across all levels of intelligence. There was no indication that approaches operated differentially for pupils with high or low intelligence.

(2) Progams which were superior in the various Basal versus Nonbasal comparisons tended to be superior across all levels of auditory discrimination ability. There was no indication that approaches operated differentially for pupils with high and low auditory discrimination.

(3) Programs which were superior in the various Basal versus Nonbasal comparisons tended to be superior across all levels of pre-instructional letter knowledge. There was no indication that approaches operated differentially for pupils with high or low ability to recognize letters.

(4) There is no basis for using test information of the nature employed in this analysis to place pupils differentially in a Basal program or any other instructional program utilized in this investigation. A teacher who is successful with a given instructional program will probably be successful with that approach for pupils of varying degrees of readiness and capability.

Conclusions from the Basal versus Language Experience comparison are less clearcut. There is some indication that low readiness pupils are more successful in a Basal program, while high readiness pupils profit more from a Language Experience program. However, this trend must be studied further in light of the fact that the finding in this investigation on which the conclusion is based may have resulted from sampling problems.

APPENDIX I

Individual studies of Cooperative Research Program in First-Grade Reading Instruction

An evaluation of three approaches to teaching reading in first grade (Project 2719). Director: Elizabeth Ann Bordeaux, Goldsboro City Schools, Goldsboro, North Carolina.

Identified the most effective of three methods of teaching reading to first-grade children. Twenty-seven first-grade classes in the Goldsboro City Schools were divided into three groups of 9 each. All three groups used the North Carolina basal text program—the Scott-Foresman series (Robinson et al., 1960–62). Group A, considered the control group, used only materials being used at the time the study began. Group B, in addition to the basal text program, used an intensive phonetic approach. Group C used both the basal program and the intensive phonetic approach in addition to a sensory experience approach.

A study in depth of first-grade reading (Project 2728). Director: Jeanne S. Chall, The City College of the City University of New York.

Investigated the effect of interactions on various components of reading achievement between: (1) the published reading program, (2) the teacher's implementation and understanding of that program, and (3) the varying characteristics of the pupils.

Children from 12 first-grade classes in socially disadvantaged neighborhoods in New York City were subjects. Fourteen teachers participated, including two teachers who replaced two others who left during the study. The teachers were chosen for the sample had indicated their beliefs and practices concerning the teaching of reading in the first grade on a questionnaire given them prior to the study. The sample of teachers chosen for the study represented equal numbers of meaning and sound-symbol emphasis teachers as well as experienced and inexperienced teachers within each emphasis. All of the teachers followed the reading programs they had used in previous years, which were eclectic basal reader approaches.

Four teachers were observed once a week; the other eight were observed once a month for the eight months' period. Ratings of teacher characteristics and practices in teaching reading were made for each observed lesson, using a Classroom Observation Inventory constructed for the study. In addition, an interview was conducted with each teacher to obtain more information about reading practices and procedures.

The relationships of the children's initial skills and abilities, the professed methods used, and the teacher's implementation of those methods to the final reading achievement measures were analyzed.

Comparison of the basal and the coordinated language-experience approaches in first-grade reading instruction (Project 2729). Director: Donald L. Cleland, University of Pittsburgh.

Determined the effects and outcomes of teaching beginning reading to superior pupils from three levels of social strata by two different methods. Superior pupils were assigned to 24 classrooms; 12 classes used the basal reader approach to first-grade reading instruction, and 12 classes used the coordinated language-experience approach.

Supplementary materials to enrich the program for superior pupils were used in the group using the basal reader approach. The coordinated language-experience approach emphasized oral expression of ideas and utilized the stories told by the children, retaining as nearly as possible the language patterns of the children. Later in the program, self-selection of reading materials was permitted and use was made of teacher-made worksheets and programed self-corrective type materials for reinforcement of needed skills.

First grade reading instruction using diacritical marking system, initial teaching alphabet, and basal reading system (Project 2745). Director: Edward B. Fry, Rutgers–The State University.

Compared three methods of beginning reading instruction using 21 first-grade classrooms from three middle class suburban school districts in central New Jersey. Two of the methods under investigation were a diacritical marking system, developed by the principal investigator, and the Initial Teaching Alphabet—writing systems which offered greater regularity than the traditional set of basic reading texts.

The materials used for the i.t.a. group were the *Early To Read Series* (Tanyzer & Mazurkiewicz, 1964). The diacritical marking system classes used the Sheldon Readers (Sheldon et al., 1957) with diacritical marks superimposed on the words. The traditional set of basic reading texts used was the Sheldon Readers.

A study of the relative effectiveness of three methods of teaching reading in Grade One (Project 2687). Director: Harry T. Hahn, Oakland Schools, Pontiac, Michigan.

Tested the effectiveness of three approaches to teaching first-grade reading: the language arts approach, the Initial Teaching Alphabet, and the basic reader approach. In the 12 school districts, one classroom was assigned to each of the three approaches. Thus, the study

(continued)

comprised 36 classrooms in which children were matched on the basis of performance demonstrated in kindergarten as well as on socioeconomic status.

The language arts approach encouraged individual expression through a variety of media. After a firm language-experience relationship was established, a balance of directed group reading and individualized reading was included. The i.t.a. approach employed materials prepared for schools in England plus some structured materials prepared from Initial Teaching Alphabet Publications, Inc. The basic reader approach used controlled vocabulary and systematic instruction procedures in basic reading texts and workbooks normally found in a first-grade classroom.

Comparing reading approaches in first-grade teaching with disadvantaged children (the CRAFT project) (Project 2677). Investigators: Albert J. Harris and Blanche L. Serwer, The Research Foundation of the City University of New York.

Compared the relative effectiveness of two major approaches to teaching reading to disadvantaged urban children: (1) the skills-centered approach, and (2) the language-experience approach. Each of these was tried with two variations, making four treatment methods in all. These four treatment methods were: (1) a skills-centered method using basal readers, with close adherence to the instructions contained in the teacher's manuals; (2) a skills-centered method utilizing basal readers, but substituting the phonovisual method of teaching word-attack skills for the word-atttack lessons accompanying the basal reader; (3) a language-experience method, in which the beginning reading materials were developed from the oral language of the children; and (4) a language-experience method with heavy supplementation of audiovisual procedures.

Twelve elementary schools, each with a very high percentage of Negro children and a minimum of six first-grade classes, were selected for the study. There was a random assignment of the four methods to schools, two methods to each school.

An attempt to secure additional evidence concerning factors affecting learning to read (Project 2697). Director: Robert B. Hayes, Department of Public Instruction, Harrisburg, Pennsylvania.

Refined, extended, and strengthened knowledge of beginning reading by comparing methods and materials in four approaches. The four programs and the materials used were: (1) an eclectic, "whole word" reading program as represented by the Scott-Foresman Company (Robinson et al., 1960); (2) a "phonic" reading program as represented by the J.B. Lippincott Company (McCracken & Walcutt, 1963); (3) a combination eclectic, "whole word-phonic" reading program as represented by the Scott-Foresman materials (Robinson et al., 1960), supplemented with the Phonics and Word Power (Johnson et al., 1964); (4) a language arts approach using the Initial Teaching Alphabet as a medium, represented by the i/t/a Publications Inc. (Tanyzer & Mazurkiewicz, 1963). Ten elementary schools and 20 first grades were selected for this study.

Effects of an intensive in-service program on teacher's classroom behavior and pupil's reading achievement (Project 2709). Director: Arthur W. Heilman, The Pennsylvania State University.

Studied the effects of an intensive in-service program on (1) teacher's classroom behavior and (2) reading achievement of pupils taught by participating teachers.

Thirty first-grade teachers of Williamsport, Pennsylvania, public schools volunteered for the experiment. Half of the group was selected at random to serve as the experimental group and the remaining group of teachers served as the control group.

The teachers in the experimental group (1) attended and participated in a two-week preschool seminar and (2) attended and participated in 25 two-hour seminar sessions held during the first 30 weeks of the school year. The preschool seminar was devoted to examining research and the implications of research for first-grade teachers. The weekly meetings were devoted to sharing teaching techniques and diagnostic procedures.

A comparison between the effect of intensive oral-aural Spanish language instruction and intensive oral-aural English language instruction on the reading readiness of Spanish-speaking school beginners in grade one (Project 2648). Director: Thomas D. Horn, The University of Texas.

Tested the hypothesis of no difference among the effects of three kinds of oral language instruction on the reading readiness of Spanish-speaking grade 1 pupils. The treatment groups were (1) oral-aural English intensive language instruction, (2) oral-aural Spanish intensive language instruction, and (3) no intensive oral-aural language instruction.

Twenty-eight classes were arbitrarily assigned to one of the three treatments: nine to oral-aural English, ten to oral-aural Spanish, and nine to no oral-aural treatment.

(continued)

APPENDIX I (cont'd.)

Individual studies of Cooperative Research Program in First-Grade Reading Instruction

The first method involved intensive oral-aural English instruction 1 hour each day—30 minutes by demonstrator and 30 minutes by teacher. The second method concentrated on oral-aural Spanish intensive instruction 1 hour per day with the same division of time. Each of these instruction methods was used in place of the usual 1 hour prereadiness instruction. The third group, considered the control group, received no intensive oral-aural instruction.

A comparative study of two first grade language arts programs (Project 2576). Director: William M. Kendrick, Department of Education, San Diego County, San Diego, California.

Determined the relative effectiveness of the experience approach to the teaching of the language arts, as compared with the traditional method. Four areas of language arts were separately measured: reading, writing, listening, and speaking. In addition, an index of development in reading interest was taken and pupil attitude toward reading was determined.

The experience approach used the language and thinking of individual children as the basis for skill development. The traditional method group adhered very closely to the teacher's manual for each reader in the Ginn Series (Ousley & Russell, 1964) as a guide to instructional procedures. Fifty-four teachers, 27 for each treatment group, participated in the study. The pupil population of the study came from 43 elementary schools of 17 school districts located in various parts of San Diego County.

An experimental study of the group versus one-to-one instructional relationship in first grade basal reading programs (Project 2674). Director: James B. MacDonald, University of Wisconsin.

Compared the effects of ability grouping with a one-to-one instructional relationship in beginning reading instruction. Seventeen classrooms, seven experimental and ten control, were involved in the project.

After the usual readiness program was completed, one group instituted a one-to-one relationship, while the other used ability grouping. Both groups employed typical basal materials.

Evaluation of level designed visual-auditory and related writing methods of reading instruction in grade one (Project 2650). Director: John C. Manning, Fresno State College.

Compared the effectiveness of materials and techniques which were programed at various ability levels on pupil reading achievement in grade 1. Thirty-six classes in all were utilized with 13, 12, and 11 classrooms in treatment groups I, II, and III, respectively.

In the first treatment group, the teacher's manual accompanying the Ginn Basic Reading Series (Ousley & Russell, 1964) was used to develop the instructional materials. With the second group, basic visual and auditory discrimination skills in letter knowledge, word recognition, word meaning, and word analysis were stressed and subsequent reading instruction was programed in a levels design using the Ginn Series for vocabulary and story content only. In addition to the basic reading program used with the second group, written language procedures were used with the third group. A 10-level design allowing for maximum learning rate differences was followed in the latter group.

A comparative study of reading achievement under three types of reading systems at the first grade level (Project 2659). Director: Sister M. Marita, Marquette University.

Compared (1) a basal approach, using three to five groups within a class; (2) an individualized approach in which sight vocabulary is built through experience charts and reading proceeds through self-selection of books and individual conferences with the teacher; and (3) an experimental approach which was a modification and combination of the language-experience and the basal approaches. In the experimental approach, individual differences were provided for through independent reading, more intensive instruction when needed, and other enrichment activities. The sample was composed of 30 classes from the Milwaukee suburban public schools. Ten classes were used for each of the three systems under investigation.

First grade reading using modified co-basal versus the initial teaching alphabet (Project 2676). Director: Albert J. Mazurkiewicz, Lehigh University.

Compared reading achievement of two matched groups at the end of first grade. Both groups used the language arts approach: one used co-basal materials printed in traditional orthography, while the other used the Initial Teaching Alphabet materials. The study included 30 first-grade classrooms divided into two groups of 15 classes, each matched on the basis of intelligence.

The hypothesis tested was that method, rather than medium, is responsible for the differences in reading achievement and that, if method is controlled, no significant differences in reading achievement would be found.

(continued)

APPENDIX I (cont'd.)

Individual studies of Cooperative Research Program in First-Grade Reading Instruction

A study of approaches to first grade English reading instruction for children from Spanish-speaking homes (Project 2734). Director: Roy McCanne, Colorado State Department of Education.

Tested the hypothesis that there is no difference in achievement in reading English in first grade between pupils who speak Spanish at home and are taught by a conventional English readiness and basal reader approach, those who are taught by a modified TESL (Teaching English as a Second Language) approach, and those who are taught by a language-experience approach. The project also provided and organized data to aid in determining a specific sequence of skills that is appropriate for first-grade children from Spanish-speaking homes who are learning to read in English, and identified appropriate materials and techniques for teaching these skills in a culturally integrated first-grade classroom.

Subjects in this study were nonmigrant first-grade children. Their classrooms were culturally integrated, containing 12 to 20 children from Spanish-speaking homes plus children from English-speaking homes. Total class sizes ranged from 25 to 30 pupils.

A study of two methods of reading supervision (Project 2706). Director: Katherine A. Morrill, University of Hartford.

Determined the feasibility of improving the reading achievement of first-grade children by a change in the role of the reading consultant in work with teachers. Two methods of consultant help were used. One method was that of a typical consultant role on a one-to-one basis in which the consultant served teachers on request from the teachers directly or from the building principal. The other method was that of a consultant role designed to foster teacher interaction. In this method, the consultant brought together teachers with the common problem of first-grade reading instruction to share methods, materials, procedures, problems, and ideas in scheduled meetings on released time. It was hoped that the interaction would result in more knowledgeable and more skilled teachers as evidenced by the greater achievement of their pupils and that the study would show that a consultant can serve several teachers at a time in a limited number of sessions, thus increasing her effectiveness beyond that when she works on a one-to-one basis.

The total first-grade population of 10 elementary schools in Wallingford, Connecticut, comprising 35 first-grade classrooms with a like number of teachers, was utilized in the study. Seventeen teachers were exposed to the usual consultant procedure, and 18 teachers were released for one-half day twice a month for a series of meetings with the reading consultant and the other teachers in this group.

Reading achievement in relation to growth in perception of word elements in three types of beginning reading instruction (Project 2675). Director: Helen A. Murphy, Boston University.

Examined (1) the relation of perception of word elements to sight vocabulary growth, (2) the effect of early teaching of a speech-based phonics program on reading achievement, and (3) the value of a writing emphasis in the speech-based phonics program.

Three different reading programs were included in the study; each was used in 10 first-grade classrooms. One group followed the "gradual phonics approach" found in the Scott-Foresman (Robinson et al., 1960–62) readers and workbooks. A second group followed the systematic Speech-to-Print Phonics with a visual word study. The third group also used the Speech-to-Print Phonics with an emphasis on writing responses.

The 30 classrooms involved were located in three industrial cities. Five classrooms from each of two communities comprised the population for Treatment A; 5 other classrooms from each of the same two communities comprised Treatment B; and 10 classrooms from a third community furnished the population for Treatment C. Care was taken to include at least three classrooms in each treatment group in "culturally deprived" areas.

Evaluation of three methods of teaching first grade reading to children likely to have difficulty with reading (Project 2702). Director: Olive S. Niles, Springfield Public Schools, Massachusetts Department of Education, Boston.

Determined whether first-grade children who have been identified by a series of tests as likely to have greater than usual problems in learning to read could be helped most effectively by (1) using the regular basal program which is used by all other children in their classroom, (2) using the regular basal program together with remedial teacher time assigned to serve the class of which they are a part, (3) using materials other than the regular basal program which is used by the other children in the class, or (4) using a combination of remedial teacher time and materials other than the regular basal program.

One group had a supplementary remedial teacher. The remedial teacher worked with the regular classroom teacher, giving special attention to children in the poten-

(continued)

tial problem group. Regular basal readers were used. Another group was provided with special materials for the potential problem group. The children were given thorough instruction with a set of readiness materials. When they achieved success with these, they were put into library-type or trade books rather than basal readers. The third group was provided with both the additional teacher time and the use of the special materials. The fourth group was the control group. No changes were made in procedures and the regular basal program was used.

The effect of different approaches of initial instruction on the reading achievement of a selected group of first-grade children (Project 2698). Director: Hale C. Reid, Cedar Rapids Public Schools, State University of Iowa.

Compared seven methods of teaching reading to the low reading group in 45 classrooms. In each classroom, an average of eight pupils were in the lowest reading group. The seven methods were: (1) a language method involving reading, writing, listening, and speaking; (2) a method involving recognition of letters and their sounds and the use of context clues; (3) a functional approach built around easy-to-read books; (4) Skills Development Method; (5) a combination of Method 1, language, and Method 2, letter sounds; (6) a combination of Method 1, langauge, and Method 3, literature; (7) a combination of Method 1, language, and Method 4, Skills Development.

The effect of four programs of reading instruction with varying emphasis on the regularity of grapheme-phoneme correspondences and the relation of language structure to meaning on achievement in first grade reading (Project 2699). Director: Rober B. Ruddell, University of California, Berkeley.

Investigated the effect on word recognition and reading comprehension of published and specially prepared reading programs varying in the degree of regularity of grapheme-phoneme correspondences programed into the vocabulary presented and the emphasis on language structure as related to meaning.

Pupils in 24 classrooms took part in the study of four reading programs: (1) a program which used a basal reading series with little provision for emphasis on language structure as related to meaning; (2) a program which used a set of programed reading materials with vocabulary utilizing consistent grapheme-phoneme correspondences to a high degree, but placing little emphasis on language structure as related to meaning; (3) a program which used a basal reading series (same as 1 above) supplemented by materials designed to build an awareness and understanding of language structure as related to meaning; and (4) a program which used a set of programed reading materials (same as 2 above) supplemented by materials designed to build an awareness and understanding of language structure as related to meaning.

A secondary consideration of the investigation was the study of the relation of selected language and background variables to reading achievement in each of the four programs.

Comparison of reading achievement of first grade children taught by a linguistic approach and a basal reader approach (Project 2666). Director: J. Wesley Schneyer, University of Pennsylvania.

Compared the reading achievement of first-grade children taught by the Fries linguistic approach with that of children taught by a basal reader approach. Each group consisted of 12 classes: four of above average, four of average, and four of below average intelligence levels.

The two methods differ in the amount of emphasis given to word discrimination and word meaning. The linguistic approach places emphasis upon the word discrimination principle, which is based upon a mastery of sound-symbol relationships of spoken language as expressed in spelling patterns. The objective of this approach is to develop an automatic response and a rapid recognition on the part of the reader to the words in various major spelling patterns. Irregular or nonpatterned words are learned as sight words.

The basal reader places heavy initial emphasis upon meaning. Attention is focused upon regularity of the meaning-frequency-repetition principle, rather than upon regularity of the sound-symbol relationship.

Effect of first-grade instruction using basal readers, modified linguistic materials and linguistic readers (Project 2683). Director: William D. Sheldon, Syracuse University.

Compared the reading achievements of children taught by three methods of instruction. Twenty-one classrooms were divided among the three methods.

One group used a basal reading program, concentrating on direct small group instruction on children's ability levels at a rate commensurate with their ability to learn. Another group used modified linguistic instruction consisting of materials published by the Singer Company (Stern et al., 1963). The series of books progresses in difficulty so that it is possible for teachers to group children for instruction. The third group used the linguistic approach consisting of the Barnhart-Bloomfield Linguistic

(continued)

APPENDIX I (cont'd.)

Individual studies of Cooperative Research Program in First-Grade Reading Instruction

Readers (Bloomfield & Barnhart, 1963). A library of 100 easy-to-read books was installed within each classroom and children were given the opportunity to practice their reading skills using these materials for 30 minutes each day. The lowest third of each class was presented listening-viewing activities with equipment from a center consisting of a tape recorder, a record player, and a filmstrip projector.

A study of a longitudinal first grade reading readiness program (Project 2742). Director: George D. Spache, Florida State Department of Education, Tallahassee.

Determined the effect of an intensified and extended reading readiness program upon first-grade reading achievement. The "intensified and extended readiness program" consisted of a plan of instruction which utilized materials that would theoretically contribute to the development of auditory discrimination, visual discrimination, and auditory language ability, and which delayed the introduction into formal reading of pupils in the second, third, and fourth quarters of the readiness achievement distributions for periods of approximately 2, 4, and 6 months, respectively.

The design of the study provided for the inclusion of all first-grade pupils in two schools (one white and one Negro) in each of eight Florida county school systems. Of these, the eight schools in four counties served as experimental schools and the eight schools in the other four counties served as comparison schools and were designated as control schools.

Individualized reading versus a basal reader program at first grade level in rural communities (Project 2673). Director: Doris U. Spencer, Johnson State College, Vermont.

Compared the effectiveness of an individualized reading method designed to meet the needs and challenge the abilities of first-grade pupils with the basal reader method. Twenty-two teachers were selected on the basis of supervisor's ratings, interest in the project, education, and experience to participate in the project. Twelve elected to teach by the individualized plan and 10 chose to follow the Scott-Foresman Basal Reader program (Robinson et al., 1960–62).

The individualized method was based on the premise that the reading program becomes more effective as individual needs are determined and instruction is concentrated at points of weakness. The instructional program was divided into two parts: an intensive systematic phonetic instruction and a motivated varied program of story reading. This method differs from the popular concept of individualized reading as a program of self-selected story reading unsupported by systematic instruction on word skills and comprehension.

Effectiveness of a language arts and basic reader approach to first grade reading (Project 2679). Director: Russell G. Stauffer, University of Delaware.

Compared the effects of a language arts approach and a basic reader approach to teaching reading. The language arts approach utilized the children's oral language facility to develop an initial reading vocabulary and initial word attack skills, as well as group-type reading instruction in basic readers and individualized reading instruction using trade books. The basic reader approach utilized basic readers, skill books, and teachers' manuals designed to develop and maintain a reading vocabulary and word attack skills.

The sample was comprised of 20 first-grade classrooms: 10 used the language arts approach; 10 used the basic readers.

A comparison of the effectiveness of three different basal reading systems on the reading achievements of first grade children (Project 2720). Director: Harold J. Tanyzer, Hofstra University.

Compared the effectiveness of three basal reading systems: (1) a basal series with intensive emphasis upon phonics, (2) a basal reading program by Mazurkiewicz and Tanyzer utilizing the Initial Teaching Alphabet, and (3) a regular basal reading series which utilizes an eclectic approach. The study included 26 classrooms from three school districts on Long Island, New York. The children were divided not only by sex, but also by intelligence, to determine whether any of the basal systems have a differential effect, and whether they prove more successful with males than females or more successful with children of high, average, or low intelligence.

Reading achievements of first grade boys versus first grade girls using two approaches: a linguistic approach and a basal reader approach with boys' and girls' groups separately (Project 2735). Director: Nita M. Wyatt, University of Kansas.

Investigated (1) whether first-grade boys would make greater gains in reading achievement through the use of materials based on a linguistic approach than they would through the use of basal readers based on the frequency of word usage, (2) whether first-grade boys would make greater gains if they were grouped on the basis of sex and ability rather than if they were grouped on ability alone with no regard for sex, and (3) whether

(continued)

APPENDIX I (cont'd.)

Individual studies of Cooperative Research Program in First-Grade Reading Instruction

girls would make greater gains under each of the approaches studied than would boys.

Two experimental groups and one control group, each consisting of 10 first-grade classes, were organized. Children from 3 elementary school districts were chosen to constitute the sample of 633 subjects.

In one experimental group, children in 10 classes were grouped by sex as well as by ability for reading instruction. In this group, bright boys used the Houghton-Mifflin basal readers (McKee et al., 1966), while other boys used either Houghton-Mifflin or Ginn (Ousley & Russell, 1964) readers. Girls read the Scott-Foresman (Robinson et al., 1962) materials and any other supplementary materials available except those published by Houghton-Mifflin or Ginn.

With the children in the second group of 10 classrooms, a linguistic approach to reading was used. The basal program consisted of Book I of the Royal Road Readers (Diack & Daniels, 1960) published by Chatto and Windus of London, the preprimers of the Harper & Row Linguistic-Science Readers (Stratemeyer & Smith, 1963), and the Primer and Level 1–1 and 1–2 books of the Basic Reading Series (McCracken & Walcutt, 1963) published by the Lippincott Company.

The third group used materials published by Scott-Foresman, Ginn, and Houghton-Mifflin.

APPENDIX II

Instructional materials

Betts, E.A., & Welch, Carolyn M. *Betts basal readers.* (3rd ed.) New York: American Book Company, 1963.

Bloomfield, L., & Barnhart, C.L. *Let's read.* Detroit: Wayne State University Press, 1961.

Buchanan, Cynthia D. *Programmed reading.* New York: McGraw-Hill, 1963.

Bumpass, Fay L. *We learn English series.* New York: American Book Company, 1963.

Diack, H., & Daniels, J.C. *The royal road readers.* London: Chatto & Windus, 1960.

Downing, J.A. *Downing readers.* London: Initial Teaching Publishing Company, 1963.

Fries, C.A., et al. *A basic reading series developed upon linguistic principles.* Columbus, Ohio: Merrill, 1963–65.

Johnson, Eleanor, M., Singleton, C., & Wonsavage, Elaine. *Phonics and word power.* Program I. Columbus: American Education Publications, 1964.

McCracken, G., & Walcutt, C.C. *Basic reading.* New York: J.B. Lippincott, 1963.

McKee, P., et al. *Reading for meaning.* (4th Ed.) Boston: Houghton-Mifflin, 1966.

Murphy, Helen, & Durrell, D. *Speech to print phonics.* New York: Harcourt, Brace and World, 1965.

Ousley, Odille, & Russell, D.H. *Ginn basic readers.* (Revised ed.) Boston: Ginn & Co., 1964.

Robinson, Helen M., Monroe, Marion, Artley, A.S., & Greet, W.C. *The new basic readers.* Chicago: Scott-Foresman, 1960–62.

Sheldon, W.D., et al. *Sheldon basic readers.* New York: Allyn & Bacon, 1957.

Stern, Catherine, et al. *Structural reading series.* Syracuse, New York: L.W. Singer, 1963.

Stratemeyer, Clara, & Smith, H.L., Jr. *The linguistic science readers.* Evanston, Illinois: Harper and Row, 1963.

Tanyzer, H.J., & Mazurkiewicz, A.J. *Early to read i.t.a. program.* New York: i/t/a Publications, 1964.

APPENDIX III

Tests

Intelligence Test

Pintner-Cunningham primary test, form A. (General ability tests, revised). R. Pintner, et al. New York: Harcourt, Brace and World, 1946.

Readiness Tests

Metropolitan readiness tests. Gertrude Hildreth, et al. New York: Harcourt, Brace and World, 1964.

Murphy-Durrell diagnostic reading readiness test. H. Murphy & D. Durrell. New York: Harcourt, Brace and World, 1964.

Thurstone pattern copying test. Thelma G. Thurstone. *Unpublished.* University of North Carolina.

Identical forms test. L.L. Thurstone & T.E. Jeffrey. Chapel Hill: Psychometric Laboratory, University of North Carolina, 1956.

Reading Achievement Tests

Detroit word recognition test. Eliza F. Ogelsby. New York: Harcourt, Brace and World, 1953.

Gates word pronunciation test. A.I. Gates. Adapted for the U.S. Office of Education Studies from the Gates Reading Diagnostic Tests. New York: Bureau of Publications, Teachers College, Columbia University, 1942.

Gilmore oral reading test, form A. J.U. Gilmore. New York: Harcourt, Brace and World, 1951.

Phonetically regular words oral reading test. E.B. Fry. Designed for the U.S. Office of Education Studies. New Brunswick, N.J.: Rutgers University, 1964.

Stanford achievement test, primary I level, form X. K.L. Truman, et al. New York: Harcourt, Brace and World, 1964.

Reading Inventories

An inventory of reading attitudes. Monograph, no. 4. San Diego County. Department of Education, Reading Study Project Committee, 1961.

Teacher inventory of approaches to the teaching of reading. Monograph, no. 3. San Diego County. Department of Education, Reading Study Project Committee, 1961.

REFERENCES

AGNEW, D.C. *The effect of varied amount of phonetic training on primary reading.* Duke University Research Studies in Education, no. 5. Durham, North Carolina: Duke University Press, 1939.

ALLEN, R.V. The language experience approach. In W.G. Cutts (Ed.) *Teaching young children to read.* Washington, D.C.: Office of Education, U.S. Department of Health, Education and Welfare, 1964. 59–67.

ALLEN, R.V., & SAN DIEGO PUBLIC SCHOOLS. Three approaches to teaching reading. In *Challenge and experiment in reading. Proceedings of the International Reading Association,* 1962, *7,* 153–154.

ANDERSON, I.H., et al. Age of learning to read and its relation to sex, intelligence and reading achievement in the sixth grade. *Journal of Educational Research,* 1956, *49,* 447–453. (a)

ANDERSON, I.H. The relationship between reading achievement and the method of teaching reading. *University of Michigan School of Education Bulletin,* 1956, *27,* 104–108. (b)

AUSTIN, MARY C., & MORRISON, C. *The first r: The Harvard report on reading in elementary schools.* New York: Macmillan, 1963.

BALOW, I.H. Sex difference in first grade reading. *Elementary English,* 1963, *40,* 303–306.

BEAR, D.E. Phonics for first grade: A comparison of two methods. *Elementary School Journal,* 1959, *59,* 394–402.

BEAR, D.E. Two methods of teaching phonics: A longitudinal study. *Elementary School Journal,* 1964, *64,* 273–279.

BETTS, E.A., & WELCH, CAROLYN M. *Betts basal readers.* (3rd ed.) New York: American Book Company, 1963.

BLIESMER, E.P., & YARBOROUGH, BETTY H. A comparison of ten different beginning reading programs in first grade. *Phi Delta Kappan,* 1965, *46,* 500–504.

BLOOMER, R. An investigation of an experimental first-grade phonics program. *Journal of Educational Research,* 1960, *53,* 188–193.

BLOOMFIELD, L., & BARNHART, C.L. *Let's read.* Bronxville, N.Y.: C.L. Barnhart, Inc., 1963.

BOND, G.L., & DYKSTRA, R. *Coordinating Center for first-grade reading instruction programs.* (Final Report of Project No. X-001, Contract No. OE-5-10-264) Minneapolis: University of Minnesota, 1967.

BUCHANAN, CYNTHIA D. *Programmed reading.* New York: McGraw Hill, 1963.

BUMPASS, FAY L. *We learn English series.* New York: American Book Company, 1963.

BUSWELL, G.T. *Fundamental reading habits: A study of their development.* Chicago: University of Chicago Press, 1922.

CARLINE, D.E. An investigation of individualized reading and basal text reading through pupil achievement and teacher performance. Unpublished doctoral dissertation, Pennsylvania State University, 1960.

CARROLL, MARJORIE W. Sex differences in reading readiness at the first grade level. *Elementary English,* 1948, *25,* 370–375.

CHASNOFF, R.E. *Comparison of the initial teaching alphabet with the traditional alphabet in first grade reading.* (Report of Cooperative Research project no. S-210) Washington, D.C.: U.S. Office of Education, 1965.

CLYMER, T., & ROBINSON, HELEN M. Reading. *Review of Educational Research,* 1961, *31,* 130–144.

DAVIS, D.C. Phonemic structural approach to initial reading instruction. *Elementary English,* 1964, *41,* 218–223.

DIACK, H., & DANIELS, J.C. *The royal road readers.* London: Chatto & Windus, 1960.

DOLAN, SISTER MARY EDWARD. A modified linguistic versus a composite basal reading program. *The Reading Teacher,* 1964, *17,* 511–515.

DOLCH, E.W., & BLOOMSTER, M. Phonic readiness. *Elementary School Journal,* 1937, *38,* 201–205.

DOWNING, J.A. *Downing readers.* London: Initial Teaching Publishing Company, 1963.

DOWNING, J.A. *Experiments with Pitman's initial teaching alphabet in British schools.* London: Initial Teaching Publishing Company, 1963.

DOWNING, J.A. *The initial teaching alphabet.* New York: Macmillan, 1964.

DOWNING, J.A. Teaching reading with i.t.a. in Britain. *Phi Delta Kappan,* 1964, *45,* 322–329.

FIDELIA, SISTER MARY. The relative effectiveness of Bloomfield's linguistic approach to word attack as compared with *Phonics we use.* Unpublished doctoral dissertation, University of Ottawa, 1959.

FRIES, C.C. et al. *A basic reading series developed upon linguistic principles.* Columbus, Ohio: Merrill, 1963–65.

FRIES, C.C. *Linguistics and reading.* New York: Holt, Rinehart & Winston, 1962.

FRY, E.B. *Phonetically regular words oral reading test.* (Designed for the U.S. Office of Education Studies.) New Brunswick, N.J.: Rutgers University, 1964.

GATES, A.I. *Gates word pronunciation test.* (Adapted for the U.S. Office of Education Studies from the Gates Reading Diagnostic Tests.) New York: Bureau of Publications, Teachers College, Columbia University, 1942.

GATES, A.I. Results of teaching a system of phonics. *The Reading Teacher,* 1961, *14,* 248–252. (a)

GATES, A.I. Sex differences in reading ability. *Elementary School Journal,* 1961, *61,* 431–434. (b)

GATES, A.I., & RUSSELL, D.H. Types of materials, vocabulary burden, word analysis, and other factors in beginning reading. I and II. *Elementary School Journal,* 1938, *39,* 119–128.

GATES, A.I., et al. A modern systematic versus an opportunistic method of teaching: An experimental study. *Teachers College Record,* 1926, *27,* 679–700.

GIBBONS, HELEN D. Reading and sentence elements. *Elementary English Review,* 1941, *18,* 42–46.

GILMORE, J.V. *Gilmore oral reading test, form A.* New York: Harcourt, Brace and World, 1951.

GOLDBERG, LYNN, & RASMUSSEN, D. Linguistics and reading. *Elementary English,* 1963, *40,* 242–247.

GRAY, W.S. The teaching of reading. In *Encyclopedia of Educational Research.* New York: Macmillan, 1960. Pp. 1114–1135.

GROFF, P.J. Materials for individualized reading. *Elementary English,* 1961, *38,* 1–7.

GUNDERSON, DORIS V. *Research in reading at the primary level.* (Final report of contract OE-30008) Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1963.

GUNDERSON, DORIS V. *Research in reading readiness.* (Final report of contract OE-30013) Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1964.

HARRINGTON, SISTER MARY J., & DURRELL, D.D. Mental maturity versus perception abilities in primary reading. *Journal of Educational Psychology,* 1955, *46,* 375–380.

HARRIS, T.L. Summary of investigations relating to reading. *Journal of Educational Research,* 1964, *57,* 283–336.

HEILMAN, A.W. *Teaching reading.* Columbus, Ohio: Merrill, 1961.

HILDRETH, GERTRUDE, et al. *Metropolitan readiness test.* New York: Harcourt, Brace and World, 1964.

HOLMES, J.A., & SINGER, H. Theoretical models and trends toward more basic research in reading. *Review of Educational Research,* 1964, *34,* 127–150.

JOHNSON, ELEANOR M., SINGLETON, C., & WONSAVAGE, ELAINE. *Phonics and word power.* Program I. Columbus: American Education Publications, 1964.

BEST COPY AVAILABLE

LOBAN, W.D. *The language of elementary school children.* Champaign, Illinois: National Council of Teachers of English, 1963.

MAC KINNON, A.R. *How do children learn to read.* Vancouver: Copp Clark, 1959.

MAZURKIEWICZ, A.J. Teaching reading in America using the initial teaching alphabet. *Elementary English,* 1964, *41,* 766–772.

MC CRACKEN, G., & WALCUTT, C.C. *Basic reading.* New York: J.B. Lippincott, 1963.

MC CRISTY, ANTOINETTE. A comparative study to determine whether self-selective reading can be successfully used at second grade. Unpublished master's thesis, University of Southern California at Los Angeles, 1957.

MC DOWELL, REV. J.B. Report on the phonic method of teaching children to read. *Catholic Education Review,* 1953, *51,* 506–519.

MC KEE, P., et al. *Reading for meaning.* (4th ed.) Boston: Houghton-Mifflin, 1966.

MORGAN, E.F., JR., & LIGHT, M. A statistical evaluation of two programs of reading instruction. *Journal of Educational Research,* 1963, *57,* 99–101.

MORRONE, V.E. A critical analysis of scientific research in phonics. Unpublished doctoral dissertation, University of Pittsburgh, 1958.

MURPHY, HELEN, & DURRELL, D. *Murphy-Durrell diagnostic reading readiness test.* New York: Harcourt, Brace and World, 1964.

MURPHY, HELEN, & DURRELL, D. *Speech to print phonics.* New York: Harcourt, Brace and World, 1965.

NEW YORK, THE UNIVERSITY OF THE STATE OF. THE STATE EDUCATION DEPARTMENT, DIVISION OF RESEARCH. *A critical review of research and opinion on phonics.* New York: Author, 1963.

OGELSBY, ELIZA F. *Detroit word recognition test.* New York: Harcourt, Brace and World, 1953.

OLSON, A.V. Growth in word perception abilities as it relates to success in beginning reading. *Journal of Education,* 1958, *140,* 25–36.

OUSLEY, ODILLE, & RUSSELL, D.H. *Ginn basic readers.* (Revised ed.) Boston: Ginn & Co., 1964.

PAULEY, F.R. Sex differences and legal school entrance age. *Journal of Educational Research,* 1951, *45,* 1–9.

PINTNER, R., et al. *Pintner-Cunningham primary test, form A.* (General ability tests, revised.) New York: Harcourt, Brace and World, 1946.

PRESCOTT, G.A. Sex differences in Metropolitan Readiness Test results. *Journal of Educational Research,* 1955, *48,* 605–610.

PRESTON, R.C. Reading achievement of German and American children. *School and Society,* 1962, *90,* 350–354.

ROBINSON, HELEN M. Summary of investigations relating to reading. *The Reading Teacher,* 1964, *17,* 326–392.

ROBINSON, HELEN M., MONROE, MARION, ARTLEY, A.S., & GREET, W.C. *The new basic readers.* Chicago: Scott-Foresman, 1960–62.

RUDDELL, R.B. An investigation of the effect of the similarity of oral and written patterns of language structure on reading comprehension. Unpublished doctoral dissertation, Indiana University, 1963.

RUDISILL, MABEL. Interrelations of functional phonic knowledge, reading, spelling and mental age. *Elementary School Journal,* 1957, *57,* 264–267.

RUSSELL, D.H., & FEA, H. Research on teaching reading. In N.L. Gage (Ed.) *Handbook of research on teaching.* Chicago: Rand, McNally, 1963. Pp. 865–928.

SAFFORD, A.L. Evaluation of an individualized reading program. *The Reading Teacher,* 1960, *13,* 266–270.

SAN DIEGO COUNTY, DEPARTMENT OF EDUCATION, READING STUDY PROJECT COMMITTEE. *An inventory of reading attitudes.* Monograph, no. 4. San Diego: Author, 1961.

SAN DIEGO COUNTY, DEPARTMENT OF EDUCATION, READING STUDY PROJECT COMMITTEE. *Teacher inventory of approaches to the teaching of reading.* Monograph, no. 3. San Diego: Author, 1961.

SARTAIN, H.W. The Roseville experiment with individualized reading. *The Reading Teacher,* 1960, *13,* 277–281.

SHANE H.G., & MULRY, JUNE G. *Improving language arts instruction through research.* Washington, D.C.: Association for Supervision and Curriculum Development, NEA, 1963.

SHELDON, W.D., et al. *Sheldon basic readers.* New York: Allyn & Bacon, 1957.

SPARKS, P.E., & FAY, L.C. An evaluation of two methods of teaching reading. *Elementary School Journal,* 1957, *57,* 386–390.

SPERRY, FLORENCE. What research says about individualized reading. *Proceedings of the Claremont Reading Conferences,* 1961, *16.*

STERN, CATHERINE, et al. *Structural reading series.* Syracuse, New York: L.W. Singer, 1963.

STEWART, D.K. Values and limitations of basal readers. *Materials for Reading,* 1957, no. 86.

STEWART, REBECCA W. i.t.a.—after two years. *Elementary English,* 1965, *42,* 660–665.

STRATEMEYER, CLARA, & SMITH, H.L., JR. *The linguistic science readers.* Evanston, Illinois: Harper and Row, 1963.

STRICKLAND, RUTH G. Implications of research in linguistics for elementary teaching. *Elementary English,* 1963, *40,* 168–171.

STRICKLAND, RUTH G. The language of elementary school children: Its relationship to the language of reading textbooks and the quality of reading selected children. *Bulletin of the School of Education, Indiana University,* 1962, *38*(4).

STROUD, J.B., & LINDQUIST, E.F. Sex differences in achievement in the elementary and secondary schools. *Journal of Educational Psychology,* 1942, *33,* 657–667.

TANYZER, H.J., & MAZURKIEWICZ, A.J. *Early to read i.t.a. program.* New York: i.t.a Publications, 1964.

TATE, H. The influence of phonics on silent reading in grade I. *Elementary School Journal,* 1937, *37,* 752–763.

TEMPLIN, MILDRED C. *Certain language skills in children.* Minneapolis: University of Minnesota Press, 1957.

TENSUAN, E., & DAVIS, F. Phonic method vs. combination method in teaching beginning reading. *Reading Newsletter,* 1963.

THURSTONE, L.L., & JEFFREY, T.E. *Identical forms test.* Chapel Hill: Psychometric Laboratory, University of North Carolina, 1956.

THURSTONE, THELMA G. Thurstone pattern copying test. *Unpublished manuscript,* University of North Carolina.

TIFFIN, J., & MCKINNIS, M. Phonic ability: Its measurement and relation to reading ability. *School and Society,* 1940, *51,* 190–192.

TRUMAN, K.L., et al. *Stanford achievement test: primary I level, form X.* New York: Harcourt, Brace and World, 1964.

WAETJEN, W.B., & BRAMBS, JEAN D. Sex differences: A case of educational evasion? *Teachers College Record,* 1963, 65, 261–271.

ZIRBES, LAURA, et al. *Practice exercises and checks on silent reading in the primary grades.* New York: Columbia University, 1925.

Reprinted from Reading Research Quarterly
Vol. 32, No. 4
October/November/December 1997
©International Reading Association

P. David Pearson

The First-Grade Studies: A personal reflection

It is a special privilege for me to comment on the First-Grade Studies during our commemoration of the 30th anniversary of their publication. They have a special place in my memory and my personal history, for it was during the early years of my graduate study at the University of Minnesota that Bond and Dykstra were completing the final stages of data analysis and manuscript development for the Cooperative Research Program in First-Grade Reading.

Remembrances of things past

Several images of those times and the people associated with the Studies leap readily to mind.

- Most vivid is the image of my fellow graduate student, John Litcher (now in the School of Education at Wake Forest University), trudging across the wintry Minnesota landscape with, quite literally, boxes of punch cards to feed into our state-of-the-art CDC computer. That computer, which occupied a space the size of two ordinary college classrooms, served the computing needs of the entire research community on campus with roughly the computing power of the typical 486 machine that now sits atop the computer table in our offices. Somehow, we managed, but not without many trips across that wintry landscape.
- Another image—the day the first batch of final reports (on which the 1967 *RRQ* article was based) arrived at Bob Dykstra's office. What was so special was the sense of excitement that we all felt as graduate students—as if we were in possession of privileged information, as if we had the inside scoop on a major professional secret. And we did, for a few days at least. It was also the first occasion on which I met Guy Bond (who, I learned later, had reviewed my application to graduate school).
- Graduate seminars in which we examined the First-Grade Studies in juxtaposition with Jeanne Chall's newly released book, *Learning to Read: The Great Debate* (1967), and some of the interpretive pieces written in the wake of the momentous report. Of particular interest among us was the question of whether the Studies did show that code-emphasis approaches were superior to meaning-emphasis approaches in promoting reading achievement.
- Frequent discussions with John Manning, my advisor and a principal investigator for one of the 27 individual first-grade studies, about the nature of beginning reading instruction, the early reading program that he had put together in Clovis, California, and how the statistical practice of analysis of covariance had obscured the positive impact of his kindergarten intervention program on first-grade reading achievement. You see, a major component of John's intervention was an intensive readiness program in kindergarten in which kids learned letter names and sounds as a part of a larger program to promote attention and persistence, the assumption being that if students can acquire these dispositions, they can learn in large classrooms in which group instruction requires vicarious learning. Of course, this treatment effect for students in his experimental group was wiped out when the analysis of covariance was conducted. The data on p. 65 of the original text/p. 39 in this volume corroborate John's concern.

Two different worlds

Enough of reminiscence. My task in this retrospective is to examine the First-Grade Studies from the lens of literacy research at the *fin de siècle*, and to ask what they can teach us today. To examine what they might teach us, we must first acknowledge just how different our world of reading research is now from the world in which Guy Bond and Robert Dykstra analyzed the data and wrote the text for the First-Grade Studies. We were on the cusp of several revolutions. Dick Venezky had just shown us that English orthography, when examined from a nuanced linguistic (morphophonemic) perspective, was uncommonly predictable. Ken Goodman was putting the final editorial touches on *Linguistic Cues and Miscues in Reading* as he began his quest to convince us all that reading is fundamentally a language process. Bob Ruddell and Harry Singer were probably planning the 1968 institute that would lead to the first publication of *Theoretical Models and Processes of Reading*. Psycholinguistics was emerging as a field, as was sociolinguistics, through the work of scholars such as William Labov, Joan Baratz, and Roger Shuy, who were using it as a lens for examining the role of dialect in learning to read. We were still nearly a decade away from the cognitive revolution. Postmodernism, constructivism, feminism, and critical theory, while surely alive and well in well-tended intellectual plots, were still almost 2 decades away from exerting a major influence on mainstream thinking about reading processes or practices.

Despite these rumblings from our sibling disciplines, it was fair to conclude, in 1967, that researchers and practitioners viewed reading as fundamentally a perceptual process. Whether we supported phonics, linguistic approaches, initial teaching alphabet (i.t.a.), or look-say as the most appropriate approach to beginning reading, the common view of most (by no means all) educators was that the job in reading was to turn symbols into sounds in order to get the words right and, in so doing, receive the meaning set down by the author. I belabor this point about theoretical contexts and trends so that modern readers of this remarkable inquiry can judge its worth as an intellectual endeavor from the perspective of the theoretical constructs that prevailed in the late 1960s.

The legacy of the First-Grade Studies

So what can we learn from this 30-year-old document? How does it speak to issues, concerns, and questions faced by the current generation of reading educators?

Learning about the efficacy of early reading instruction

If the only document you read is the 1967 *RRQ* version of the First-Grade Studies, the most plausible conclusion about the most effective approach to teaching beginning reading is, "It all depends." Across sites and projects, the variability is remarkable; as Bond and Dykstra conclude, "Reading programs are not equally effective in all situations." Statistically, this means that there were a fair number of project-by-treatment interactions: A treatment that rose to the top in one project (project equals site) may have achieved poor results in a second.

The second conclusion that you (as well as Bond and Dykstra) would draw is that pretty much any alternative (basal plus phonics, phonics-first, linguistic, special orthography, or language experience) was, on balance, superior to the whipping boy of the First-Grade Studies, the conventional look-say basals popular in the early 1960s. In fact, Bond and Dykstra use this finding to support two interesting conclusions: (a) that combination approaches are superior to single approaches, and (b) that reading instruction is amenable to improvement (apparently on the assumption that basals represent the conventional wisdom that stands in need of improvement).

You would not conclude from the 1967 *RRQ* report that code-emphasis approaches were superior to meaning-emphasis approaches. Even though both the phonics plus basal and the phonics first (dubbed phonic-linguistic) approaches consistently elicited higher comprehension and word reading scores than basal approaches, Bond and Dykstra avoided the conclusion. In a later article in which he compared Chall's (1967) conclusions with the span of first- and second-grade results from the First-Grade Studies, Dykstra (1968a) concluded, "Data from the Cooperative Research Program in First-Grade Reading Instruction tend to support Chall's conclusion that code-emphasis programs produce better overall primary grade reading and spelling achievement than meaning-emphasis programs" (p. 21).

One of the most consistent findings in the Studies (also replicated in the second-grade follow-up) is the singular absence of aptitude-by-treatment interactions. With the exception of the Language Experience analysis (in which, depending on the site, LEA proved uniquely suitable to either high- or low-aptitude students), there were no indications that methods were consistently effective with subgroups of students identifiable by various aptitude indices (e.g., intelligence, letter-name knowledge, or phoneme knowledge). The commonsense homily about low-aptitude students needing the structure and guidance of an explicit phonics-first approach received no support from any of these analyses. Where phonics worked well, it worked well for all students; conversely, when it worked poorly, it worked poorly for all.

Learning about contextual variables

As a part of the data-collection process, information was gathered about teacher and classroom variables (teacher experience, class size, etc.). While few of these variables showed much relationship to student achievement (they tended to correlate in the .2 to .4 range), another contextual variable, project, emerged as a powerful factor in explaining student differences. While project was mentioned in the reports of both the first- and second-grade analyses, it was extensively analyzed only in the full version of the second-grade extension of the studies, in which Dykstra (1967) compared the means for each instance of a method across projects. A common finding was that the mean for the lowest performing method in Project X was often as high as or even higher than the highest performing method in Project Y. Even controlling for a wide range of individual difference variables, there was apparently something about a particular project that elicited high (or low) performance for all students regardless of method. This sort of finding led Dykstra to conclude that future research needed to focus on site-based variation in order to learn what it was about the *culture* of a site that led to such consistently exceptional performance. His recommendation presages our current preoccupation for examining instruction and learning with a *situated* lens.

Learning about the practice of research

There is a great deal to learn about the practical concerns of conducting research, especially in settings in which many individuals are involved.

1. *Other things being equal, include multiple measures of any important phenomenon.* Clearly the principal investigators of the 27 individual studies thought it important to include multiple measures, although they have many more measures of word reading than of comprehension. It should also be mentioned that writing samples were included in the original design, but were not included in the final analysis (apparently out of concern for standardization of administration and concerns about scoring reliability). The exclusion of writing samples was a real disadvantage for the i.t.a. treatments. One of the serendipitous findings of some of the specific i.t.a. analysis was that when kids were equipped with a transparent orthography that was completely under their control, they became fearless writers, producing a great deal of text. One is reminded of the remarkable fluency of students in today's classrooms when they are encouraged to use invented spellings (e.g., Clarke, 1989).

2. *Sometimes finding an appropriate control group requires a bit of imagination.* Given the variety of interests brought to the table by the 27 principal investigators, it is truly amazing that they were able to settle on any common notion of a control group. The idea of using a *garden variety* basal, which according to research of the time was used by 95% of all schools in the United States, was a brilliant compromise. Granted, it did not allow for comparisons among the various innovative procedures, but it clearly allowed for a common benchmark and for some interesting cross-site analysis. The premise that all basals are created equal (and therefore constitute a common basis for comparison) does require a stretch but, given what we know of basal production at the time, it is probably not a far-fetched assumption.

3. *The appropriate unit for statistical analysis depends upon the question one wants to answer.* For the correlational analysis, Bond and Dykstra used the scores of individual students as the basic unit of analysis, a practice that fits the predictive nature of the question (predictions are about individuals, not classes). But for the comparison of methods and the method by treatment interaction analyses, class means (actually the separate means for boys and girls—so that gender could be included in the analysis) were used. Again, they fit the question and the situation; the method was applied simultaneously to all members of a class.

4. *Novelty can affect findings.* Both the classic Hawthorne effect (all the groups do better because they are in a study) and the more selective novelty effect (the experimental groups do better because they get privileged treatment in comparison to controls) can compromise findings in experimental studies. While attempts were made to control for novelty (e.g., making sure that even the lowly basal teachers received an equal amount of inservice training), it is hard to imagine the same feeling of newness and excitement among teachers who are part of a business-as-usual treatment. While Bond and Dykstra did not mention this possibility, other commentators on First-Grade Studies did (e.g., Sipay, 1968; see also Southgate, 1966).

5. *Treatment fidelity is a common problem in large-scale research.* In the limitations section of the second-grade extension (1968b), Dykstra explicitly mentioned the problem of treatment fidelity across sites: What was called a linguistic approach in Project X might be quite different from a linguistic approach in Project Y. (Curiously, there is no limitations section in the *RRQ* version of the First-Grade summary.) Equally problematic was ensuring the fidelity of treatments within projects. While every attempt was made to ensure fidelity, teachers occasionally applied their own standards to a method in order to make it fit their own philosophy and professional practice. (Dykstra [personal communication, 1968] tells a story about visiting a remote Language Experience site as a part of his monitoring role and finding the teacher engaged in a lesson from a commercial phonics workbook. When queried about the practice, she replied, "Well, this Language stuff is very interesting and the students really like it, but, you know, they need their phonics...." I am reminded of the incredible variation in what goes on in the name of whole

language in today's instructional milieu. I am also reminded of the highly eclectic tendencies of the teachers nominated as outstanding in Pressley's 1996 work on exemplary teachers.) Of course, the larger the study, the greater the threats to fidelity. (An alternative argument is that if effects survive the probable treatment infidelity that is likely to occur in large-scale instructional work, then they must be truly robust!)

6. *Occasionally technical standards must be compromised for the sake of credibility and utility.* In several situations, such compromises can be seen in the First-Grade Studies, but they are not bothersome. To the contrary, they permit more careful and considered inspection of findings across the different analyses. For example, in the guidelines for analysis (pp. 47–49/p. 28), a logic is put forward that IF and ONLY IF project-by-treatment interactions emerge will separate analyses be conducted for treatments within each of the projects. In truth, the within-projects analysis was conducted regardless of whether project-by-treatment interactions appeared. In another example, parsimony would dictate that a single model for analysis of covariance be selected (e.g., either all of the covariates or only those most likely to control for relevant pre-experimental variation). As it turned out, both a full and a minimal set of covariates were used in all of the analyses. For the reader who wishes to make some of his or her own eyeball comparisons, it is most useful to have the full set of analyses, quite irrespective of whether they are technically appropriate.

Learning some lessons to guide our future research

A common, but usually implicit, standard for evaluating the legacy of a piece of research is whether it generates additional studies on the issue, topic, or question. By that standard, the First-Grade Studies were a dismal failure, for they (in conjunction with Chall's book) marked the end of methodological comparisons in research on beginning reading (at least until the 1990s). Dykstra (1968b) recognized and championed this uncommon legacy:

> One of the most important implications of this study is that future research should center on teacher and learning situation characteristics rather than method and materials. The extensive range of classrooms within any given method points out the importance of elements in the learning situation over and above the materials employed.... The elements of the learning situation attributable to teachers, classrooms, schools, and school systems are obviously extremely important. Reading instruction is more likely to improve as a result of improved selection and training of teachers, improved in-service training programs, and improved school learning climates, rather than from minor changes in instructional materials. (p. 66)

By the way, in the aggregate, the principal investigators (see the Appendix of the reprint of the First-Grade Studies in this volume) involved in the larger set of First-Grade Studies understood the need to move beyond the racehorse mentality. Chall, for example, conducted an intensive study of the instructional practices used by committed code-emphasis and meaning-emphasis teachers (foreshadowing our current emphases on understanding effective practice). Heileman studied the impact of alternative models of inservice training (also foreshadowing our modern commitment to the idea that the payoff for teacher learning is student learning), and Morrill studied the impact of one-on-one versus group (learning community?) approaches to teacher supervision. And Horn anticipated our current debates over bilingual programs by comparing intensive aural-oral programs in English versus Spanish for students speaking Spanish as a first language.

As it turned out, Bond and Dykstra were prophetic in suggesting an end to methods research. By the early 1970s, we had declared a moratorium on racehorse studies, we had begun the process of changing basals in keeping with the conclusions reached by Jeanne Chall (1967) (and supported, at least in the minds of many, by the First-Grade Studies), and we turned our intellectual attention to new issues, new perspectives, and new disciplines—psycholinguistics, cognitive science, theories and models of the reading process—that lay in wait just beyond the horizon, poised to capture our hearts and minds as a profession. Only recently, driven by politics and alarmist interpretations of test scores and fueled by new sources of funding from outside the educational research industry (Lyon & Chhabra, 1996), have we returned to the question of the best method of teaching beginning reading. The fact that the most recent request for proposals for a national reading center requires an emphasis on early reading indicates that we have pretty much come full circle, back to the issues and questions that prompted us as a profession to undertake the First-Grade Studies some 35 years ago. But, as I indicated at the outset, we are at a very different time and place than we were when all of this began. We have new tools and new lenses for asking and answering questions of teaching and learning.

Even though our world on the cusp of the 21st century is very different, I do hope that we can recover one of the most endearing and important qualities of the era—the model that guided the cooperative part of the endeavor. I think it is a good model for literacy research in a postmodern world.

Like most models, it has a historical antecedent. When I first encountered the First-Grade Studies, I was immediately reminded of a metaphor stolen from that decidedly premodern thinker, the renown British romanticist, philosopher, designer, and social reformer of the late Victorian era, William Morris. As a metaphor for his ideal society, Morris chose the Gothic cathedral; as a contrastive metaphor for his evil society, the neoclassicist cathedral.

What Morris liked about the Gothic cathedral had more to do with the process of construction than the outcome, although the two are linked. To him, it represented the proper balance of order and individual freedom. There was a master planner with a master plan, to be sure, but that person was more of a foreman than an architect. Each worker was assigned responsibility for completing a particular section of the cathedral. There was enough coordination between sections to ensure the structural integrity of the edifice, but no more. Within individual sections, each worker exercised a great deal of individual prerogative. As a result, Gothic cathedrals lack the unity and precision of their neoclassicist counterparts. The design of the gargoyles in one corner is quite different from those in another. The carvings, even the stained glass, in one section may or may not match those in another. In every nook and cranny, there is the distinct mark of the individual craftsman. There was, to use modern terms, ownership and empowerment.

By contrast, the unity and precision of the neoclassicist design was mirrored in its construction. Workers carried out orders. They implemented the plans of others. There was no room for the individual signature of each worker. There were no sections individual workers to point to as theirs. So, said Morris, and partially in response to growing Marxist sentiment in England and Europe, let's build a society of Gothic cathedrals as a way of ensuring that human beings are connected to, rather than alienated from, their work.

I like that metaphor a great deal for thinking about our research, both in its micro- and macroscopic aspects. Within a project, when a group of us work together—teachers, administrators, and researchers—we need to find ways for each of us, as individuals, to put our own personal stamp on the project. There must be room for variation. But just as surely as there is variation, there must also be theme, a common core to which we are all committed. We should all be learning something different about the same thing. We can move that model one level up and think about implementing cross-site studies, studies in which teams decide to pool their intellectual and material resources to gain variable insights on questions of common interest.

This is, I believe, the model and the metaphor that guided, knowingly or unknowingly, the First-Grade Studies. Twenty-seven individual researchers or groups, each with his, her, or their own agenda, had a unique piece of the Gothic cathedral of reading research to shape in a unique image. Each, however, ceded some independence to be a part of a larger effort, to answer some bigger questions, than a single study could answer. I suppose that this sort of collaboration came with the territory so to speak, since the First-Grade Studies were funded by the Cooperative Research Branch of the then United States Office of Education.

It is a model, however, that I think we should strive to emulate, to reincarnate, both in spirit and in form. We need both theme and variation in our work. The themes bring us together, encourage us to share and collaborate on a common vision, while the variations remind us to respect, enjoy, take pride in, and, most important, learn from our differences. And to think that our intellectual predecessors were smart enough to figure that out 30 years ago! My only question is why it seems so hard for us to emulate such a sensible practice in our current world of research.

REFERENCES

CHALL, J. (1967). *Learning to read: The great debate.* New York: McGraw-Hill.

CLARKE, L.K. (1988). Invented versus traditional spelling in first graders' writing: Effects on learning to spell and read. *Research in the Teaching of English, 22,* 281–309.

DYKSTRA, R. (1967). *Continuation of the Coordinating Center for first-grade reading instruction programs* (Report of Project No. 6-1651). Minneapolis, MN: University of Minnesota.

DYKSTRA, R. (1968a). The effectiveness of code- and meaning-emphasis beginning reading programs. *The Reading Teacher, 22,* 17–23.

DYKSTRA, R. (1968b). Summary of the second-grade phase of the Cooperative Research Program in primary reading instruction. *Reading Research Quarterly, 4,* 49–70.

LYON, G.R., & CHHABRA, V. (1996). The current state of science and the future of specific reading disability. *Mental Retardation and Developmental Disabilities Research Reviews, 2,* 2–9.

PRESSLEY, M., RANKIN, J., & YOKAI, L. (1996). A survey of the instructional practices of primary teachers nominated as effective in promoting literacy. *Elementary School Journal, 96,* 363–384.

SIPAY, E. (1968). Interpreting the USOE cooperative reading studies. *The Reading Teacher, 22,* 10–16.

SOUTHGATE, V. (1966). Approaching i.t.a. results with caution. *Reading Research Quarterly, 1,* 35–56.

BEST COPY AVAILABLE

Reprinted from Reading Research Quarterly
Vol. 32, No. 4
October November December 1997
©International Reading Association

Lyndon W. Searfoss

Connecting the past with the present: The legacy and spirit of the First-Grade Studies

> Every year hundreds of thousands of children begin the complex task of learning to read. For most children, growth in reading is a successful undertaking. For many, however, the progress is slow, and for others learning to read appears to be an unobtainable accomplishment. There is a continuous search for new ways to teach reading which will prevent the difficulties these children encounter, thereby enabling all children to become successful readers.
>
> In recent years, many new approaches to reading instruction have been suggested. Many questions have been raised about current methods of teaching reading. In fact, the teaching of beginning reading has been and continues to be a popular subject for debate among reading experts and the general public alike. (Bond & Dykstra, p. 9 in original text/p. 9 in this book)

These sentences were written 30 years ago in the opening paragraphs of the 1967 issue of the *Reading Research Quarterly* devoted to the final report of the Cooperative Research Program in First-Grade Reading Instruction, known more popularly as the First-Grade Studies. These words could have been written today as the controversy surrounding how to teach beginning reading, always just beneath the surface, once again erupts in the U.S. Is it worth looking back at this 30-year-old report for insight or perspective on today's debates and discussions? I think so. Recently, Moore, Monaghan, and Hartman (1997) decried our inability as a profession to be better keepers of the past, our literacy history, as they call it. While history cannot always be instructive or inform current practices in education, Moore et al. presented some significant reasons why a look back at our literacy history is worth the time and the effort. They say that knowing and revisiting our history, in part, helps us understand the conceptual background of current issues and situations, and contributes to a "distanced view, a measured grasp of the big picture, that is difficult to attain during contemporary calls for action" (p. 92), and helps develop a shared sense of our intellectual roots.

The sheer complexity of the First-Grade Studies was lost on most of us who read it as graduate students in 1967. We thought it a simple and logical idea to take methods, teachers, and children, assign them to treatments, pretest and posttest, and then report what happened. Well, as we know, it wasn't that simple. The size of the project made it the largest of its kind to date, involving cooperation in research design across 27 individual projects. Because each individual project or study had its own director and combination of methods to study, the diversity across projects was staggering by any standard.

Of the 27 original projects, 15 were part of the final analysis and report by Guy L. Bond and Robert Dykstra. It is worth reading about some of the projects not included in the final analysis, too. For example, Horn's Texas project and McCane's Colorado study (see Bond & Dykstra, 1967, Appendix I) foreshadowed significant issues for researchers and teachers of bilingual students today because they included Spanish-speaking children. A number of projects addressed inner city and culturally deprived students (see Harris & Serwer, Chall, and especially Helen Murphy's well-designed and executed project in Appendix I).

In this brief look at the First-Grade Studies 30 years later, I know readers familiar with the studies will find issues missing from this commentary. This is intentional; other writers can, and should, debate the flaws and merits of these methods comparison studies, the research and statistical procedures used in the individual projects, the objectivity of the project directors, and the very nature of the conclusions drawn by the project directors. The focus in this commentary will be on whether these studies have led us anywhere as researchers and teachers since they were published in 1967.

In the Introduction to Bond and Dykstra's report, Theodore Clymer and Edward G. Summers cautioned readers that "a quick reading or a few quotes pulled from the summary sections will not do justice to the content of the report or the spirit of the inquiry" (p. 3/p. 7). As a classroom teacher and graduate student in 1967, I eagerly read the report for its *content* of these studies, as I searched for an answer to the question, "What is the best way to teach beginning reading?" *Spirit* was not even a consideration, especially for a classroom teacher turned graduate student at Syracuse University who wanted to read about the latest and best research on the issue. I reread the report, this time in 1997, as a way of capturing the spirit of these studies. With 30 years' distance, the names of the individual 27 project directors are really a hall of fame of reading educators and researchers, many of them now gone from the scene. For me and my fellow graduate students of the 1960s, fresh from classroom teaching, they were our teachers and mentors. We looked to them to help us find answers to important questions about how children learn to read and why some become successful readers and others struggle and struggle with the process.

Choosing a graduate program for many of us in the 1960s was made a bit easier because of the First-Grade Studies. If an institution was the site of one of the individual studies included in the final project report, it was on the top of the list of potential schools for graduate study. The mantle of a First-Grade Study was a valuable recruitment tool for colleges and universities. Syracuse University became a natural choice for me because it was an individual study site directed by William D. Sheldon, who became my doctoral advisor and mentor.

Throughout my graduate studies, the names of other individual project directors came alive as they visited our campus. Knowing some of them firsthand was one of the most exciting and rewarding aspects of graduate study. Taking courses seemed secondary at times to the opportunities we had for conversations and discussions with what seemed (at least to graduate students) these larger-than-life personalities. You took Sheldon for a class, marveled at Russell Stauffer's command of the English language during one of his visits, and were challenged by Helen Murphy's fire when she spoke of beginning readers (and learned so much from listening in on her conversations with Donald Durrell). We watched remedial readers who were tested and tutored in our university reading clinic read in i.t.a. and wondered, just wondered, whether it would really work with these children who were failing at learning to read. We took dictation for language experience stories and worried if we should correct the children's oral dictation errors as we wrote down their stories. Even the machines of O.K. Moore and his talking typewriter caught our attention, along with *Sullivan Programmed Readers* and *Words in Color* with its colorized vowel sounds to help the reader crack the code.

As we careened from one method to another in the 1960s and read the methods research, searching for which one by itself or which combination of methods would work in classroom and clinic settings, the First-Grade Studies appeared. We all looked to the report of this large-scale effort to provide the definitive answer to the question, "What is the best way to teach beginning reading?" Did these studies answer this question? Not really, I think. What they did do was influence the direction of both research and practice in the coming years by pushing both researchers and teachers into taking another look at the question itself.

As I look back, today's research topics and classroom practices can be traced, in part, to the legacy of the First-Grade Studies in at least four major ways, because they (a) led us to a closer and more discriminating look at the role of the teacher and of classroom characteristics in reading instruction, (b) added strong support for the movement away from methods comparison studies, (c) provided evidence to reinforce the efforts of many researchers who challenged the traditional view of reading readiness as a gatekeeper to beginning reading instruction, and (d) lent support for the use of phonics as a part of beginning reading instruction. Each of these legacies can be better understood in the context of the original research questions posed by the project investigators for all of the individual studies. While it is always risky to attribute major changes in research and practices to a single event, the sheer size and diversity of the First-Grade Studies mandated close attention to their findings and conclusions.

Research question #1:

To what extent are various teacher, class, school, and community characteristics related to achievement in first-grade reading and spelling?

Bond and Dykstra concluded, after considerable discussion, that factors other than those examined in these studies influenced success in beginning reading.

> Reading achievement is related to other characteristics in addition to those investigated in this study. Pupils in certain school systems became better readers than pupils in other school systems even when pupil characteristics were controlled statistically. Furthermore, the differences in achievement from project to project do not seem to be directly related to the class, school, teacher, and community characteristics appraised in this study. (pp. 122–123/p. 75)

They recommended that "Future research might well center on teacher and learning situation characteristics rather than method and materials" (p. 123/p. 75). It was clear to the researchers that, frankly, few of the teacher characteristics they studied really mattered. Teacher data collected included (a) sex, (b) age, (c) degrees earned, (d) certification, (e) years of teaching experience, (f) years of experience teaching first grade, (g) marital status, (h) number of children, (i) attitude toward teaching, (j) number of days absent during study, and (k) supervisor's rating of teacher effectiveness. Some of these measures were dropped from the final data analysis because they were judged as unreliable (teacher effectiveness rating) or unrelated to reading achievement (number of days absent, score on teacher attitude scale).

Bond and Dykstra's conclusion received much attention, and the door was opened for using the First-Grade Studies as a major source of research supporting studies on the nature of effective teaching and also the role of class environments as created by effective teachers. Researchers, tentatively in the 1970s and with greater interest in the 1980s, began to conduct studies in these areas. The shift to these topics is reflected in how the First-Grade Studies were discussed in subsequent research reviews after 1967. For example, they appear in two different contexts in the two volumes of the *Handbook on Reading Research*. In the 1984 volume, the First-Grade Studies were included in a chapter on beginning reading instruction written by Rebecca Barr. By the 1991 volume, the First-Grade Studies were reported in a chapter by James Hoffman entitled, "Teacher and School Effects in Learning to Read." Hoffman commented about methods comparison studies such as the First-Grade Studies:

> Most of the methods-comparison studies conducted in the field of reading operated under the assumption that teachers, once told or trained in what to do, would do just that. Seldom were teachers actually observed in the implementation of the methods, nor was much consideration given to what the individual teacher already knew or believed about the method. It was assumed in these studies that, aside from the experimental manipulation, all teachers were doing pretty much the same thing in their classrooms. (p. 929)

The current interest and research in the dual roles a teacher can play as both a teacher and a researcher is a logical extension of this movement toward illuminating how and why teachers do what they do.

Research question #2

Which of the many approaches to initial reading instruction produces superior reading and spelling achievement at the end of first grade?

Bond and Dykstra concluded:

> No one approach is so distinctly better in all situations and respects than the others that it should be considered the one best method and the one to be used exclusively. (p. 123/p. 75)

The First-Grade Studies marked the beginning of a decline in methods studies, especially by major researchers, that continues today. These studies led us away from the inevitable frustration with methods studies in classrooms to other avenues of exploration and inquiry about beginning reading. In the 30 years since the First-Grade Studies were published, we seemed almost to transcend classroom methods studies as a profession and move to studying the processes involved in reading and toward generating theories that explained how reading occurs.

This research path included going beyond traditional, narrow topics of inquiry (vocabulary, comprehension, and word recognition, for example) to cross-disciplinary research in areas such as psycholinguistics, sociolinguistics, cognition, sociology, psychology, social constructivism, and literary criticism, to name a few. We have even tried to apply theories from other fields of inquiry that seem to many of us to be only remotely related to reading (e.g., chaos theory and Marxist viewpoints) in our efforts to better explain the reading process. From this research, we then try to arm teachers with frameworks, paradigms, and even metaphors to use as they choose methods and materials for teaching reading in their classrooms. We exhort teachers to become decision makers and reflective practitioners, based on theory and, ultimately, on their own personally constructed theories.

At this point, some small voice in my head asked about how the methods used in 1967 have fared over the years. A look at the following lists reveals what seems to be fewer choices, with commercial, published programs still in wide use, although they are generally much more broadly conceived than the basal readers of the 1960s. Phonics, regardless of what adjectives appear before or after it, is still with us, as is the term *eclectic* when applied to teachers who say they take what they need from a variety of methods.

1967 methods	*1997 methods*
Commercial basal reading series	Commercial literacy programs
Initial Teaching Alphabet	Literature-based reading instruction
Language experience	Whole language
Phonics-based	Integrated/thematic instruction
Linguistic	Skills (phonics)-based
Individualized reading	Eclectic
Eclectic	

Research question #3

Is there any program uniquely effective or ineffective for pupils with high or low readiness for reading?

The pupil data collected for the First-Grade Studies included information on sex, chronological age, preschool experience, and number of days absent during the experimental period. A battery of measures of reading readiness was also included, since it was generally agreed that children who were *ready* to read would be most successful in beginning reading instruction. But a challenge to this assumption emerged as the data were analyzed collectively and individually across projects. Differences existed across projects in pupils' reading achievement, even after initial pupil differences in readiness were controlled statistically. As Bond and Dykstra stated:

> One of the most striking findings was the persistence of project differences in reading achievement, even after adjustments were made statistically for differences in pupil readiness for reading. Evidently, reading achievement is influenced by factors peculiar to school systems over and above differences in pre-reading capabilities. (pp. 121–122/ p. 74

It was indeed significant to the project directors that they could not depend solely on readiness tests to predict success. Was it some other unknown and unmeasured factor that determined success once again coming through the numbers and statistics? Or, as some experts argued, was the traditional view of readiness in need of scrutiny? The concept of readiness for reading as a both a predictor and a determinant of later success in learning to read was dominant, and beginning reading practices adhered firmly to it in the 1960s. Formal and informal tests of readiness assessed a variety of factors deemed necessary for success in beginning reading. These tests served as a gate through which children had to pass before formal reading instruction. In fact, as some readers may recall, a score on a readiness test was often used for grouping children for instruction and for retaining some children in kindergarten. These readiness notions of the 1960s have given way to a focus on understanding the roots of children's literacy or their emerging literacy.

One of the unintended or indirect consequences of the First-Grade Studies, I believe, was the support it gave to reading educators and others who were challenging the prevailing views of reading readiness. The First-Grade Studies included widely used, traditional measures of readiness. Today, we look at the emerging reader a bit differently.

1967 readiness	*1997 emerging literacy*
Auditory discrimination	Print awareness
Recognizing upper & lower case letters	Phonemic awareness
Word learning	Concepts of print
Copying a figure	Story sense
Figure discrimination	Oral language
Vocabulary	Writing
Listening—following directions	Letter identification
Word recognition	High-frequency words

Finally, one of the last legacies of the First-Grade Studies is the widely reported conclusion that these studies helped solidify the place of phonics and letter-name knowledge in beginning reading instruction. Coming on the heels of Chall's (1967) study, they added more evidence to the historical pile of studies supporting phonics instruction that can be traced almost as far back in literacy history as you want to go. Both the First-Grade Studies and Chall's study were often oversimplified over the years as studies supporting a code emphasis or strong phonics strand to early reading instruction. It is interesting that in today's current furor over phonics, they are appearing once again as support for phonics instruction, although they offer little in the way of specifics about what kinds of phonics instruction are effective. So, for the third or fourth time in my career (you'd think we would get it right soon; how complicated can it really be?), phonics is in the spotlight again. Once again my phone rings and a reporter from the local newspaper or the campus newspaper asks for my opinion on phonics vs.__________________ (you fill in the latest iteration).

A summary and a look to the future

Did the First-Grade Studies lead researchers on a different path or paths? Did they influence classroom practices? Because of the sheer size and cooperative nature of the project, the diversity across projects, and the unheard-of amount of federal funding for the time, the First-Grade Studies commanded our attention in 1967. I like to think they did leave a legacy over the next 30 years that directly influenced research and practice in some ways (e.g., the need to explore the role of the teacher and classroom environments, the shift away from methods comparison studies) and indirectly in other ways (e.g., giving momentum to a changing view of reading readiness). They were a necessary part of helping us understand that we needed to go beyond trying to determine what way to teach reading and to probe deeper to reveal how reading occurs as a process.

Clymer and Summers told us we would find a spirit to the inquiry of the First-Grade Studies, if we did a careful reading. My own recent reading and thinking for this retrospective has led me to wonder how we can keep that spirit alive during the next generation of research in beginning reading. I also hear the voices of Moore and his colleagues (1997) reminding me how important it is to preserve and make public our literacy history. By revisiting these studies, I have come to the conclusion we need to make active inquiry into methods and materials of beginning reading a legitimate activity once again, using the tools of inquiry not available to the researchers in the First-Grade Studies and guided by questions we now can generate 30 years later. I fear we have made it passé and anti-intellectual, at times, for teachers to ask the research community to help them find out what works. Why is asking "What should I do on Monday morning?" greeted with glee by countless companies marketing workshops for teachers and with disdain (or at least mild concern that it is the wrong question) by many of us at colleges and universities? Have we been too timid to suggest that there is nothing dishonorable or anti-intellectual about methods comparisons studies? These questions now deserve our attention.

A small, strident bell has been sounded by Thomas (1997) in a recent article entitled "What's the Use of Theory?" in the *Harvard Educational Review*. It is challenging reading and should be discussed by all of us who mentor graduate students before another doctoral student is urged to ground or situate a study in theory before a dissertation proposal can be accepted. Thomas observed that in education we have difficulty defining what we mean by theory, since we borrowed that term, along with others, from the sciences, and that we cling to theory with a tenacity and persistence that stifles creativity and is hard for us to justify. He cautioned that:

> Theory's acquired potency for bestowing academic legitimacy is troublesome, for it means that particular kinds of endeavor in educational inquiry are reinforced and promulgated, while the legitimacy of atheoretical kinds is questioned or belittled. (p. 76)

Thomas made a case for what he calls "ad hocery" in educational inquiry rather than theory generation and argued that "creativity and progress are rarely the fruit of theory and more often the fruit of anarchy in thought" (p. 77). Quoting Medawar (1974), he also cautioned that educators often are seduced by theory because of poeticism, "namely the adoption of a theory because of its elegance, attractiveness, or romantic appeal" (p. 78). Theory building and studies to validate theories seem to have an intellectual air about them that methods studies do not.

Theories can be destructive, too. Thomas rightly pointed out that we clung too long to the notion of readiness for reading rather than moving on to new understandings. Shouldn't theories be viewed as a tool to understanding and encouraging creativity, rather than as something to be proven as valid and then defended even in the face of evidence that refutes or debunks? But, I do remember there still are people who believe the Earth is flat and that the Moon landings were staged by the Air Force somewhere in a hangar in New Mexico. Maybe we learn more from being atheoretical than from trying to ground our work in a theory.

The spirit of the inquiry that Clymer and Summers said could be found in reading the First-Grade Studies lives on in reading research today. In that same spirit we need to remember what we are about—helping children learn to do that grand and powerful and wonderful thing of learning to read. If the methods studies of 1967 did not answer how we should teach reading, and the subsequent years of theory generation have not been as productive as we thought, then we must move on to new ways of thinking. Perhaps we should worry less about which theory we are following and instead use the tools for inquiry that did not exist in 1967, both quantitative and qualitative, to get on with exploring and replicating why some teachers in some classrooms are highly successful in helping children learn to read. I can't think of anything more exciting than to blur the lines between research and practice, guided and inspired by the past rather than encumbered and limited by it. For here is the paradox: Our literacy history is both a blessing and a curse. It blesses us by guiding and pushing us to new ways of thinking, while at the same time it makes us fearful when we cannot relate unorthodox ways of thinking or doing to the familiar and the comfortable.

REFERENCES

BARR, R. (1984). Beginning reading instruction. In P.D. Pearson, R. Barr, M.L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 545–581). New York: Longman.

CHALL, J. (1967). *Learning to read: The great debate.* New York: McGraw-Hill.

HOFFMAN, J.V. (1991). Teacher and school effects in learning to read. In R. Barr, M.L. Kamil, P. Mosenthal, & P.D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 911–950).

MEDAWAR, P.B. (1974). *The hope of progress.* London: Wildwood House.

MOORE, D.W., MONAGHAN, E.J., & HARTMAN, D.K. (1997). Values of literacy history. *Reading Research Quarterly, 32,* 90–102.

THOMAS, G. (1997). What's the use of theory? *Harvard Educational Review, 67,* 75–104.

Reprinted from Reading Research Quarterly
Vol. 32, No. 4
October/November/December 1997
©International Reading Association

Arlette Ingram Willis
Violet J. Harris

Expanding the boundaries: A reaction to the First-Grade Studies

We like to think of history as alive, complex, and intersecting with events, people, and time. This thinking is not original: it is much like the thinking of the late Paulo Freire (Freire & Fraser, 1997) who described history as "never frozen" (p. 311), and of Edward Said (1993) who described history's power "to shape our lives" (p. xxii). In framing our response, we have decided to follow the format offered by our fellow commentators and begin by offering a brief remembrance of the period under discussion, 1955–1967, from our perspective. We cannot offer the firsthand, insider, privileged points of view comparable to those of our colleagues, but we can offer an alternative viewpoint of the same period. We were alive then, but were very young and in very different circumstances than the other commentators. As young children during this period we completed first grade and moved on through grade school. Thus our perspectives are informed by our own memories and our understanding of the significant historical, ideological, political, economic, social, cultural, and racial events of the period.

Life in the 60s

In the United States, the mid-1950s to late-1960s was a period of tremendous social unrest, racial tension, economic inequities, and political injustice. We offer examples of first-grade experiences that reveal the diversity of experiences among African Americans and help frame our discussion in the next few pages.

Arlette: I lived on Fourth Street in a mid-sized town in northeastern Ohio with my parents and two brothers. Like all the children in our school district, I walked to the neighborhood school, Jefferson Elementary School, with the other children on my street. For me, it was a time of wonder and excitement. I had a wonderful teacher, Mrs. Carter, to whom I will forever be indebted. She not only taught us to read, but she taught us to believe in ourselves as learners. We used a basal reading series, complete with its white, middle-class, nuclear family: Father, Mother, Alice, Jerry, baby Susan, and Jip.

Mrs. Carter taught using the transmission model of direct instruction. She combined the basal and phonics approaches to beginning reading. It seemed as if the entire morning of each school day was spent in a reading-related activity. I loved Mrs. Carter. She was strict, but we always believed that she cared for us as individuals. Mrs. Carter was an African American teacher. I can remember how excited my parents were to learn that I would have a black teacher, one they thought would care for me and about my education in ways that they were not certain white teachers would care.

Though Mrs. Carter seldom was absent, when she was, our substitute was Mrs. Blakely, another African American teacher. Mrs. Blakely and her family attended our church, and her daughter and I were in the same Sunday School class. When Mrs. Blakely came, I knew I had to be on my best behavior; if not, then my parents would receive a call and learn of my misbehavior before I could walk home. I understand now that Mrs. Carter and Mrs. Blakely showed their concern for my education by keeping my parents informed of my progress and behavior.

I also remember the summer of 1965, when my cousin from Montgomery, Alabama, spent her summer vacation with us. She thought it was strange that I lived near and went to school with white children. I was equally amazed that she did not go to school with white children but, instead, went to an all-black elementary school. I wondered, why should it be so unusual to go to school with the people in your neighborhood? In our town, black and white children who lived next door to one another went to the same school.

Violet: In 1962, I attended Bryant Elementary School in Chicago, Illinois. My parents viewed my education as very important and took great care with my school work. I walked to school with the other children in my neighborhood but little else stands out in my memory of first grade. The following facts will help to contextualize my experience: The students were all African American, the teacher was a Euro-American female, and the reading textbooks were basals. Chicago Public Schools were under the direction of Benjamin Willis, who acquired notoriety for the Willis Wagons. These were mobile units placed in many African American schools to accommodate excess enrollment and to avoid desegregation of the schools.

Individually, we learned that our experiences in first grade were not the same as many other African American girls. We are reminded of the story of Ruby Bridges, who became the first African American child to integrate an all-white elementary school in New Orleans in 1960. She and her family faced angry mobs daily and the white parents removed their children from her first grade classroom. For nearly a year, Ruby and her teacher experienced first grade together. Eventually, some of the white children returned as the school year continued (Coles, 1995). The horror of the time, for many African American schoolgirls, is most dramatically evinced in the 1963 Sunday morning church bombing in Birmingham, Alabama, that ended the lives of four African American girls: Addie Mae Collins, Denise McNair, Carole Robertson, and Cynthia Wesley.

Upheaval in many institutions characterized the mid-1950s to late-1960s in our country. Nightly, there were news reports of the latest unexplained murders of civil rights and voting rights workers; race riots; lynchings; and protests by numerous groups from minorities (African, Asian, Mexican, and Native Americans), to women, to senior citizens, to college students protesting U.S. involvement in the Vietnam War; and domestic injustices (from the rights of migrant workers to discrepancies in school funding). It was also a time of tremendous political involvement in domestic concerns from the United States Supreme Court ruling in the Brown v. Topeka, Kansas case for school desegregation to the War on Poverty, to the Civil Rights Act, to the Immigration Act of 1965, and to the Voting Rights Act. Two domestic social concerns seemed to override others: racial injustice and the culture of poverty (Harrington, 1962).

We begin our response by expanding the historical framework for understanding the First-Grade Studies beyond those offered in the prologue and retrospectives. Our historical reframing is necessary because it highlights the manner in which crucial elements of difference—race, ethnicity, language, and gender—were excluded from critical educational research during this period. Our historical framework is not an attempt to hold researchers to levels of enlightenment that reflect current thinking. Rather, we suggest that the inclusion of such factors could have resulted in fundamental educational, and by extension, societal changes. Then, we offer an alternative perspective to the studies. Next, we offer suggestions for how the sociocultural framework could have been expanded to be more inclusive. Finally, we propose an expanded conceptual framework for understanding the role of the First-Grade Studies in the history of reading research.

Expanding the historical framework

Historically, the linkages between literacy and schooling always have been of concern to those in positions of power and in control of education. This point is argued compellingly by historian Harvey Graff (1987), who wrote,

> From the classical era forward, leaders of polities and churches, reformers as well as conservers, have recognized the uses of literacy and schooling. Often they have perceived unbridled, untempered literacy as potentially dangerous, a threat to social order, political integration, economic productivity, and patterns of authority. Increasingly, however, they came to conclude that literacy, if provided in carefully controlled, structured, formal institutions created expressly for the purposes of education and transmission of literacy and supervised closely, could be a powerful and useful force in achieving a variety of important ends. (p. xxv)

Alternatively, those in opposition to the dominant group's use of power and their control of education have sought redress. For example, W.E.B. Du Bois conducted numerous research studies on Negro education between 1915 and 1920 while a member of the faculty at Atlanta University, and Horace Mann Bond wrote extensively of Negro education in the South in the 1930s.

We submit to you that the conception, design, and funding of the First-Grade Studies grew out of the need for political and educational leaders to address the national, ideological, historical, social, economic, cultural, and racial needs and concerns of the U.S. Generally, educational research has tended to overlook these forces and focus instead on sanitizing the period. One method for accomplishing this has been oversimplifying the impact of these issues and concentrating on the national political response to the launching of the Soviet Union's Sputnik as the primary impetus for educational funding.

Histories of reading research also tend to overly simplify the events that preceded the funding of the First-Grade Studies. By suggesting that the events merely were a response to the scientific advances in Russia and Flesch's publication of *Why Johnny Can't Read* (1955), historical accounts imply that it took nearly a decade for the impact of both of these events to be realized sufficiently enough to motivate the government to fund educational research in reading. Moreover, the traditional historical frameworks of reading ignore the dynamic of history: its intersection with politics, ideology, society, culture, and economics.

In the wake of the Cold War, Americans looked inward to domestic issues that had sparked violent and nonviolent demonstrations, protests, and riots. The national political, economic, social, and racial crises of the period challenged politicians, researchers, and educators to combat more directly poverty and address inequalities in educational opportunities and to uphold the nation's commitment to a democratic society of all Americans including African, Asian, Mexican, and Native Americans; Puerto Ricans; Appalachian whites; and people living in poverty. For example, in the aftermath of the Brown v. Topeka Board of Education court case (which legally dismantled the "separate but equal" doctrine and provided equal educational opportunities) some groups, like the National Association for the Advancement of Colored People (NAACP), questioned the ideological foundations of schooling from per pupil costs to opportunities for educational advancement.

One spin-off of the Civil Rights movement was protests/demonstrations by other underrepresented groups within the United States, who also sought equal rights, educational opportunities, and greater political and economic justice. Another spin-off was the establishment of private schools committed to offering a better education to African American children than they were experiencing in the public schools, like the Nairobi Day School in East Palo Alto, California.

It is difficult to imagine that the daily media coverage of some of the events of the Civil Rights Movement, the War on Poverty, Vietnam War, Voting Rights Act, Immigration Act, and the establishment of the National Organization for Women has been left out of the historical framing of discussions around the funding of the First-Grade Studies. Just months before the official preconference meeting of the First-Grade Studies, May 31–June 3, 1964 (but certainly following other meetings on the issue), President Lyndon Johnson appealed to Congress by addressing the social and economic disparity of our nation's poor. In a special address to Congress in March 1964, President Lyndon B. Johnson called for a national war on poverty. It was important then, as it is now, to address the needs of every American. In his 1965 State of the Union address, President Johnson stated that "for the primary and secondary school years we will aid public schools serving low-income families" (Johnson, p. 7). Later, he pledged that "new laboratories and centers will help our schools—help them lift their standards of excellence and explore new methods of teaching" (p. 7). When you review the First-Grade Studies you are left wondering, where are the issues of race and social economic status in these studies, in a period where both issues were daily before the nation?

The First-Grade Studies: An alternative perspective

An examination of the agreements among reading researchers during the spring 1964 preconference (Bond & Dykstra, 1964) reveals that reading researchers were interested in accounting for race, economic status, community organization, and income in their research study. More specifically, the recommendations concerning community, school, and class included (a) socioeconomic status of the community as derived from census reports, (b) organization of the community (urban, suburban, small town, rural), (c) occupational characteristics of the community (industrial, agrarian, etc.), and (d) general statement concerning condition of the community (stable, transitional, blighted, etc.). During the preconference it also was recommended that pupil characteristic data would include (a) sex (gender), (b) chronological age, (c) race/ethnicity, and (d) preschool experience.

The original studies' data on these variables appear to suggest that some reading researchers were aware of and concerned with the need to address the intersection of race, class, and gender and their effects on beginning reading instruction. Jean Chall (1967), director of one of the originally funded proposals, offered intriguing insights into the project. She noted that there were 76 research proposals submitted for consideration, 27 of which were funded, and 15 of which were used in the cross-study analysis. The following three questions were to be investigated in the First-Grade Studies:

1. To what extent are various pupil, teacher, class, school, and community characteristics related to pupil achievement in first-grade reading and spelling?
2. Which of the many approaches to initial reading instruction produces superior reading and spelling achievement at the end of first grade?

3. Is any program uniquely effective or ineffective for pupils with high or low readiness for reading? (p. 5 in original study/p. 9 in this volume)

In this article, we will concentrate on the impact of the first question on reading research, for it appears to indicate an interest in considering the effect of contexts upon beginning reading. Chall (1967) stated that, "The goal [of the project] was not only to compare and contrast the effectiveness of the alternative approaches to beginning reading but to identify the nature and degree of variation owed to characteristics of the child, teachers, and schools" (p. 40). Moreover, she wrote, "The question of urgency concerned...what should be expected to work best for the students in any particular classroom" (p. 40). Overall, concerns with ideological and methodological problems with the study and with the conclusions have been covered by others elsewhere (e.g., Chall, 1967; Sipay, 1968, to name a few). The problem with the use of standardized testing begs issues not addressed in this essay but addressed by others (Gould, 1981/1996; Sokal, 1987). Issues with reading tests and what they actually measure also have been well documented by Farr and Carey (1969/1986).

Several of the funded research proposals that were eliminated from analysis clearly identified students of color and those whose first language was not English, and communities that were disadvantaged. Although these were not reasons given for their exclusion from cross-study analysis, it appeared that the data from all of these studies did not meet the needs of the cross-study analysis. For example, studies not included in the cross-study analysis were directed by Horn, McCann (who conducted research among Spanish-speaking students), Chall, Harris and Sewer, and Murphy (among nonwhite populations).

One can only assume that race was negligible in the other studies because it is not mentioned as a descriptor. For example, the Coleman Report conducted in 1964 found "that almost 80 percent of all white pupils in the 1st and 12th grades attend schools that were from 90 to 100 percent white. Sixty-five percent of black students in the first grade attended schools with student populations that were between 90 to 100 percent black. This meant that the majority of children in the United States attended segregated schools and that white children were the most segregated" (Spring, 1994, p. 317). Nevertheless, the cross-study analysis included clearly identified selective sampling among white students. This selective inclusion resulted in the cross-study analysis of essentially the same type of students, white and middle-class, who were taught in selective ways (basals vs. nonbasals). Chall (1967) reported a survey by Barton and Wilder (1964) that indicated 98% of first-grade teachers used basals; thus any deviation would have some Hawthorne effect on the students, teachers, community, and research study. So, in seemingly normal classes of white or mostly white students, there was some degree of agreement among variables and the possibility for cross-study analysis.

As important as issues of race and class were in the mid-1950s to the late-1960s and in spite of their inclusion in the first question of this project, their worth to the project becomes a muffled point since the studies in which communities, schools, and the race of pupils were clearly defined (as nonwhite and non–English-speaking) were eliminated from the cross-study analysis in the Final Report. At first, it appeared odd to us that teacher race was not a question that the researchers considered. We wondered why, if pupil race was an important factor, why not the race of the teachers? From the data and reports it isn't clear why this was not a characteristic the researchers thought was important. We may never know; but here is what we do know. First, several studies have detailed, statistically, the large number of African American teachers that were displaced with the enforcement of desegregation plans during this period (Foster, 1997), thus perhaps limiting the possible pool. Second, when race isn't mentioned, it's usually because everyone is of the same race, in these cases white, though sex (gender) is an issue that was considered. The absence of race and social economic class issues, then as now, suggests deeper issues are unexamined within the field and within research. And their absence is blinding.

Ironically, several of the findings reported by Bond and Dykstra in their Final Report also reflect the need for a greater understanding of the effect of community, school, class, and teachers on beginning reading instruction. Bond and Dykstra (1967) listed the following findings:

1. Reading programs are not equally effective in all situations. Evidently, factors other than method, within a particular learning situation, influence pupil success in reading;
2. Reading achievement is related to other characteristics in addition to those investigated in this study. Pupils in certain school systems became better readers than pupils in other school systems even when pupil characteristics were controlled statistically. Furthermore, these differences do not seem to be directly related to the class, school, teacher, and community characteristics appraised in this study;
3. Future research might well center on teacher and learning situation characteristics rather than method and materials. The tremendous range among classrooms within any method points out the importance of elements in the learning situation over and above methods employed. (pp. 122–123/p. 75)

BEST COPY AVAILABLE

Bond and Dykstra (1967) declared that, "One of the most striking findings was the persistence of project differences in reading achievement even after adjustments were made for differences in pupil readiness for reading. Evidently, reading achievement is influenced by factors peculiar to school systems over and above differences in measured pre-reading capability of pupils" (p. 100/p. 62). Later, Dykstra (1968), speaking on the implications of the study for classrooms, wrote "The elements of the learning situation attributable to teachers, classrooms, schools, and school systems obviously play a larger role [in beginning reading]" (p. 11). Later he noted that, "it is likely that improvement in reading instruction can be brought about more efficiently by improved selection and training of teachers, by improved in-service training programs, and by informed school learning climates than by institutional changes in instructional programs" (p. 11).

In summary, one of the purposes of this brief discussion of the marginalization of the issues of race and class in the First-Grade Studies is to begin to examine some of the underlying assumptions that have stymied reading research. We began with an examination of the complex and interwoven relationships among historical, ideological, political, social, cultural, economic, and racial forces and events operating in the nation in the mid-1950s to late-1960s. The point we want to make here is that race, class, ethnicity, gender, language, and geographical location have been largely avoided, marginalized, or misrepresented in much reading research until the 1980s. The report itself and the events surrounding it are written in the normative, if not autonomous (Street, 1996) fashion of reading research that ignores, marginalizes, or distances itself from the contexts in which the research takes place. Street (1984/1996) described the academy's position on literacy as autonomous, claiming that within the academy, literacy has been studied as if unaffected by social, cultural, political, economic, and institutional influences and contexts.

Given the historical lack of attention in reading research to historical, ideological, political, social, cultural, economic, and racial influences and contexts, we have offered several examples that illustrate alternative perceptions of the influence of these forces on reading during the same time period. In what follows, we want to highlight briefly the resources available at the time to help inform the First-Grade Studies.

Expanding the sociocultural frameworks

As we approach the end of this century, it is helpful to review our past accomplishments in the field of reading research in light of our varying histories. Edward Said (1993) wrote, "The fact that the United States contains so many histories...is by no means to be suddenly feared since many of them were always there, and out of them an American society and politics...were in fact created" (p. xxvi). On the one hand, some fear that the plurality of histories and voices disturbs a sense of harmony and questions the status quo of reading research. On the other hand, some welcome a more honest appraisal of the past that dispels myths of normalization and perfectly crafted research and advances the notion of the need to take a proactive stance in reading research. This later stance acknowledges and invites histories and voices often silenced or marginalized, to share in the retellings of reading history. For us, this has meant reviewing this classic study and attempting to make "visible the historically and socially constructed strengths and limitations of those places and borders we inherit and which frame our discourses and social relations" (Giroux, 1997, p. 368). By way of illustration, we briefly share two examples.

How might the principal researchers have addressed issues of race? One way would have been to include in the project the many heads of colored/Negro school systems in the segregated South who attended the University of Chicago, Harvard University, and Teachers College and had access to the current research and theories available during this period. These scholars, whose research for doctoral degrees did not appear in established journals for reasons not discussed here, could have served as directors of First-Grade Studies in segregated schools. For example, Helen Adelle Whiting, director of colored/Negro education for the state of Georgia, recorded notes in which she described her efforts to apprise teachers of current ideas about reading instruction during this period.

Another way of addressing race would have been to consider information about communities and teachers. For example, some teacher histories could have illuminated or accounted for the variance in the data collected. Michele Foster's (1997) book *Black Teachers on Teaching* described the views of African American teachers on teaching over several decades. One of the interviews she conducted was with Etta Mae Marks, who began her professional teaching career in Lindale, Texas, in 1954 and was teaching during the time of First-Grade Reading Studies. Etta Mae Marks was born and raised in Lindale and taught in the consolidated school for African American students in the segregated school system of Lindale, Texas. In 1965, with the enforcement of desegregation and the threat of withholding federal funds to the district, Lindale desegregated its school system and moved Mrs. Marks to the white elementary school. She was forced to teach in an all-white school thereafter.

beginning with summer school and followed by several years as a remedial reading teacher. Finally, she was given her own classroom, but many white parents would not send their children to her for an education. Ironically, she was lucky, for she was one of the few African American teachers able to retain their jobs after desegregation in Texas (Foster, 1997).

From our perspective, race and class have always been issues in American education, as they were in the mid-1950s to late-1960s, during the period in which the First-Grade Studies were undertaken. The presentation of the First-Grade Studies as unencumbered by race and class concerns suggests some sort of scientific purity, objectivity, cultural neutrality, and imagined truth that can come from research studies that don't deal directly with the contexts in which they are situated. In this manner, the status quo is perpetuated and cultural reproduction sustained. Historian Harvey Graff (1995) put it this way: "Although seldom stated, less often contemplated present-day conceptions, arrangements, and practices of literacy as well as schooling and learning are *historically founded and grounded* [italics in original]. They are also strong and powerfully resistant to change" (p. 323). In reading, research that fails to inform itself also helps to maintain dominant ideologies and normative practices that disproportionately benefit certain groups. The price of overlooking contextual issues, especially the issues of race, class, language, and community, is without measure.

Expanding the conceptual frameworks

Unquestionably, the First-Grade Studies represent a historical breaking point for change in reading research as experts in the field undertook the monumental task of conducting a cooperative study of beginning reading instruction in the United States. Initially, its potential for informing research and practice were unfathomable. We have all learned a great deal from the studies, both from its strengths and from its limitations. In this retrospective essay, we have offered a narrative that represents an alternative way of viewing the time, events, and contexts.

What have we learned from our review of the First-Grade Studies? First, our review of this study has revealed to us that (a) the First-Grade Studies represent a particular ideological and power relationship within the academy; (b) cooperative research efforts can be a powerful force in shaping our understanding of reading research and reading instruction; (c) there is no one best way to teach beginning reading; (d) race, class, and language are important variables to consider in beginning reading instruction; and, (e) contextual considerations (community, school district, school building, classroom, teacher, and pupil) also are important variables that can add to our understanding of beginning reading instruction. The researchers connected with the First-Grade Studies acknowledged the importance of contextual features on beginning reading, within their first question of their study. Ironically, this question though not addressed in their data analysis, became the cornerstone of their argument for the need to study these very variables and how they might affect reading instruction and achievement. Prophetically, they suggested that answers to this question may hold the most promise for understanding the variance in beginning reading.

Second, our review has made clear to us that reading research has not adequately addressed the concerns raised in this first question of the First-Grade Studies, nor have we moved very far beyond the results of the study. We know that these issues are important considerations, but what specific role they play is not clearly understood. Current National Assessment of Educational Progress data suggest that the performance standards of nonwhites on standardized measures of reading are still lagging behind those of whites, regardless of the instructional methods used. Most distressingly, we have learned that it has taken nearly 20 years for reading research to evolve to the point where we take seriously the influence of race, class, and language on how students in varying communities, school districts, schools, and classrooms are instructed in beginning reading. Dudley-Marling and Murphy (1997) wrote: "Schools are also implicated in producing and reproducing inequities related to race, class, gender, and language by favoring knowledge and pedagogical practices that privilege the skills and experiences of middle- and upper-middle class students" (p. 461). Further, we have learned that contextual considerations are the responsibility of all reading researchers and should always be a part of our thinking and focus.

Finally, we have learned that reading research is a complex multilayered, dynamic process that is ever changing. Purcell-Gates (1997) put it this way: "The epistemological assumptions and presuppositions that ground 'research' have also evolved, changing and broadening from what many of us first learned and adopted, mainly from the hard sciences, two and three decades ago" (p. 283). Importantly, as we move into the next century, with projections of an increasingly diverse student population, reading research will need to extend its boundaries of inquiry to include alternative voices, perspectives, and methodologies to increase our understanding of beginning reading instruction.

In conclusion, the historical, ideological, social, cultural, economic, and racial forces that have privileged or disprivileged certain groups within the United States become our points of reference in an examination of the

BEST COPY AVAILABLE

events that were a part of the national consciousness during the years of the First-Grade Studies and the marginalization of the events in the reported data and publications. We have argued that by expanding the historical, sociocultural, and conceptual frameworks we use to understand the First-Grade Studies, we offer an alternative way of viewing the impact of the First-Grade Studies and their place in the history of reading research.

We have used examples from the literacy histories of representatives from underrepresented groups to illustrate how the myth of normalcy and the autonomy of reading research has failed to offer broad viewpoints for understanding the legacy of the First-Grade Studies. In addition, we use these examples to thwart the perpetuation of sanitized official histories of reading research and offer in their place greater contextualization and demystification of the complex and interwoven nature of reading research. We believe that mainstream and alternative histories need to be read together to offer some balance between excessively positive or negative retellings. We believe our examination of the assumptions that have upheld reading research as neutral and apolitical can help to provide a framework for conceptualizing future reading research studies.

In 1968 Howard Gardner observed, "History never looks like history when you are living through it...it always looks confusing and messy, and it feels uncomfortable" (cited in Clifford, 1988, p. 169). We agree and, along with Purcell-Gates (1997), have observed that the history of reading research appears to be at a crossroads: The field must learn from its past and more directly address issues that have been silenced, marginalized, or ignored. Our vision for the future of reading research is proactive. We would like to see more cooperative reading research efforts that include a plurality of voices and perspectives framed upon a commitment to social change and improvement.

REFERENCES

BOND, G., & DYKSTRA, R. (1964). *Conference on coordination of accepted proposals for the Cooperative Research Program in First Grade Reading Instruction.* Final Report. (Report Number CRP- F-062). Minneapolis, MN: University of Minnesota. (ERIC Document Reproduction Service No. ED 003 422)

BOND, G., & DYKSTRA, R. (1967). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly, 2,* 5–142.

CHALL, J. (1967). *Learning to read: The great debate.* New York: McGraw-Hill.

CLIFFORD, G. (1988). *A Sisyphean task: Historical perspectives on the relationship between writing and reading instruction* (Tech. Rep. No. 433). Champaign, IL: Center for the Study of Reading, University of Illinois.

COLES, R. (1995). *The story of Ruby Bridges.* New York: Scholastic.

DUDLEY-MARLING, C., & MURPHY, S. (1997). A political critique of remedial reading programs: The example of Reading Recovery. *The Reading Teacher, 50,* 460–468.

DYKSTRA, R. (1968). *Classroom implications of the First-Grade Reading Studies.* Paper presented at the annual meeting of the College Reading Association, Knoxville, TN. (ERIC Document Reproduction Service No. ED 022 626)

FARR, R., & CAREY, R. (1969/1986). *Reading: What can be measured?* Newark, DE: International Reading Association.

FLESCH, R. (1955). *Why Johnny can't read: And what you can do about it.* New York: Harper & Row.

FOSTER, M. (1997). *Black teachers on teaching.* New York: The New Press.

FREIRE, P., & FRASER, J. (Ed.). (1997). *Mentoring the mentor: A critique of dialogue with Paulo Freire.* New York: Peter Lang.

GIROUX, H. (1993). Literacy and the politics of difference. In C. Lankshear & P. McLaren (Eds.), *Critical literacy: Politics, praxis, and the postmodern* (pp. 367–377). Albany, NY: State University of New York Press.

GOULD, S. (1981/1996). *The mismeasure of man.* New York: W.W. Norton.

GRAFF, H. (1987). *The literacy myth: Cultural integration and social structure in the nineteenth century.* New Brunswick, NJ: Transaction Press.

GRAFF, H. (1995). *The labyrinths of literacy: Reflections on literacy past and present* (revised and expanded ed.). Pittsburgh, PA: University of Pittsburgh Press.

HARRINGTON, M. (1962). *The other America: Poverty in the United States.* Baltimore: Penguin.

JOHNSON, L. (1965). *Public Papers of the Presidents of the United States: Lyndon B. Johnson (Containing the public messages, speeches, and statements of the President, 1963–1964), Book I.* Washington, DC: United States Printing Office.

JOHNSON, L. (1966). *Public Papers of the Presidents of the United States: Lyndon B. Johnson (Containing the public messages, speeches, and statements of the President, 1965), Book II.* Washington, DC: United States Printing Office.

PURCELL-GATES, V. (1997). The future of research in the language arts. *Language Arts, 74,* 280–283.

SAID, E. (1993). *Culture and imperialism.* New York: Vintage Books.

SIPAY, E. (1968). Interpreting the USOE Cooperative Reading Studies. *The Reading Teacher, 22,* 10–16.

SOKAL, M. (1987). Introduction. In M. Sokal (Ed.), *Psychological testing and American society, 1890–1930* (pp. 1–20). New Brunswick, NJ: Rutgers University.

SPRING, J. (1994). *American education* (6th ed). New York: McGraw-Hill.

STREET, B. (1984/1996). *Social literacies: Critical approaches to literacy development, ethnography and education.* New York: Longman.

THE FIRST-GRADE STUDIES 30 YEARS LATER

One of the most original and influential studies in reading research was first published in 1967 in *Reading Research Quarterly*, "The Cooperative Research Program in First-Grade Reading Instruction," better known as the First-Grade Studies. Why reprint the article 30 years later? There are many important reasons:

- The importance of history to literacy researchers,
- The continued interest in beginning literacy,
- The encouragement of methodological eclecticism, and
- The need to make this study accessible to current reading educators.

Revisiting the First-Grade Studies, an offprint of the 1997 *RRQ* issue that commemorates the 30th anniversary of their publication, provides a view of the historical implications of the First-Grade Studies and how the Studies continue to provide valuable insights for those who seek answers about effective beginning reading instruction.

This book will allow literacy educators of today and those of the future to enjoy and benefit from revisiting this historical reading research.

BEST COPY AVAILABLE

INTERNATIONAL
Reading
Association

9120